United States
Environmental Protection
Agency

Air And Radiation
(6604J)

EPA 402-R-93-001C
January 1993

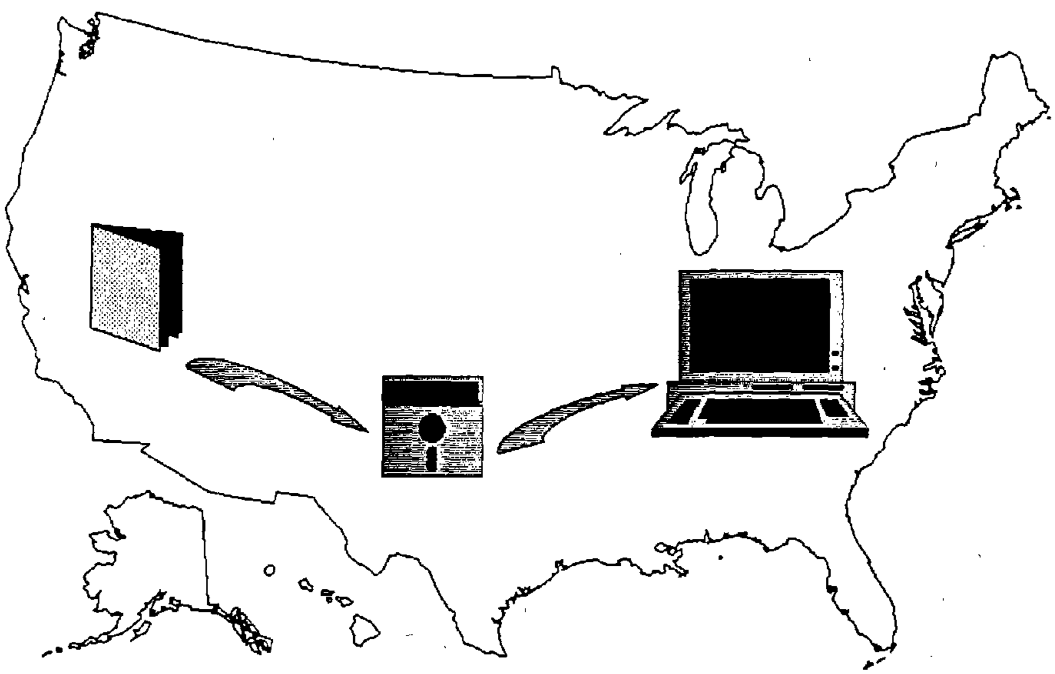
PB93-134807

EPA

National Radon Database

Volume 3: State/EPA Residential Radon Survey

PB93-134807



PB93-134807

NATIONAL RADON DATABASE DOCUMENTATION
Volume 3

The EPA/State Residential Radon Surveys: Year 3

U.S. Environmental Protection Agency
Washington, D.C. 20460

Sharon White
Work Assignment Manager

January 1993

NATIONAL RADON DATABASE
DOCUMENTATION
Volume 3

The EPA/State Residential
Radon Surveys: Year 3

Submitted by:

S. Cohen & Associates, Inc.
1355 Beverly Road
McLean, Virginia 22101

under

EPA Contract No. 68D20185
Work Assignment No. 1-3

Submitted to:

U.S. Environmental Protection Agency
Washington, D.C. 20460

Sharon White
Work Assignment Manager

January 1993

TABLE OF CONTENTS

1. INTRODUCTION 1-1
   1.1 Goals of the EPA/State Residential Radon Surveys 1-1
   1.2 Summary of the Year 3 Surveys 1-10

2. THE SAMPLE DESIGN 2-1
   2.1 The Overall Sampling Plan 2-1
   2.2 Population Definition and Sampling Frames 2-1
   2.3 Stratification and Sample Allocation 2-3
   2.4 Sample Selection Procedures for the Donnelley File Selections 2-5
   2.5 The Alaska Survey 2-7
   2.6 Partitioning the Samples into Waves 2-9
   2.7 Sample Selection Procedures for the Indian Lands Survey 2-10

3. ESTIMATION USING SURVEY RESULTS 3-1
   3.1 Calculation of Sampling Weights 3-1
   3.2 Calculating Sampling Weights for the Indian Lands Survey 3-4
   3.3 Estimating Means and Proportions 3-5

4. METHODOLOGICAL RESULTS 4-1
   4.1 Coverage 4-2
   4.2 Response Rates 4-3
   4.3 Unknown Eligibility Status 4-4
   4.4 Errors Made in Implementing Surveys 4-5

APPENDIX A INSTALLATION PROCEDURES A-1

APPENDIX B RECORD LAYOUT FOR STATE RESIDENTIAL RADON SURVEYS B-1

APPENDIX C DESCRIPTION OF SAMPLE ALLOCATION USED FOR EACH STATE C-1

APPENDIX D REGIONAL RADON COORDINATORS AND SOURCES OF INFORMATION CONCERNING OTHER STATE-WIDE RADON STUDIES D-1

APPENDIX E PROCEDURES FOR ESTIMATING WEIGHTED MEANS, PROPORTIONS, STANDARD ERRORS AND CONFIDENCE INTERVALS FOR INDIAN LANDS E-1

LIST OF TABLES

Table 1-1 Summary of Six Years of the EPA/State Residential Radon Surveys . 1-5

Table 1-2 EPA/State Residential Radon Survey Results, Years 1 to 6 1-6

Table 1-3 Parameter Estimates from the Distribution of Indoor Radon Screening Measurements in Year 3 Surveys, by State and Region (1988-89) 1-12

Table 2-1 Procedures for Selecting the Sample of Telephone Numbers 2-6

Table 4-1 Comparison of Estimates of Survey Eligibles 4-6

Table 4-2 Disposition of Sample Cases 4-7

Table C-1 Distribution of Canisters per County for Alaska C-9

Table C-1 Distribution of Canisters per County for Georgia C-10

Table C-1 Distribution of Canisters per County for Iowa C-14

Table C-1 Distribution of Canisters per County for Maine C-17

Table C-1 Distribution of Canisters per County for New Mexico C-18

Table C-1 Distribution of Canisters per County for Ohio C-19

Table C-1 Distribution of Canisters per County for Vermont C-21

Table C-1 Distribution of Canisters per County for West Virginia C-22

LIST OF FIGURES

Figure 1	Distribution of Arithmetic Means of Screening Measurements in 225 Regions	1-7
Figure 2	Distribution of Arithmetic Means of Screening Measurements in the Top 60 Regions	1-8
Figure 3	EPA Regions	1-9

1. Introduction

The National Radon Database has been developed by the U.S. Environmental Protection Agency (EPA) to distribute information collected in two recently completed radon surveys:

1. The EPA/State Residential Radon Surveys, Years 1 to 6; and
2. The National Residential Radon Survey.

The State Residential Radon Surveys were conducted in 42 states and 6 Indian lands to characterize the state-wide distribution of radon screening measurements in the lowest livable area of owner-occupied homes. The National Residential Radon Survey was designed to provide an estimate of the national frequency distribution of annual average radon concentrations in occupied residences. Data and documentation for each survey are available through the National Technical Information Service (NTIS).

1.1 GOALS OF THE EPA/STATE RESIDENTIAL RADON SURVEYS

These surveys are statistically valid at the state level and regional levels within each state. The results represent screening measurements and should not be used to estimate annual averages or health risks. Although states and portions of states have been characterized with high or low indoor radon results, the only way to determine the indoor radon level of an individual house is to test. EPA recommends that all homes test for elevated indoor radon levels.

In response to the growing concern about potential health risks associated with indoor radon exposure, the EPA initiated a program in 1986 to assist states in measuring radon concentrations in homes. The importance of this program was confirmed by the Indoor Radon Abatement Act of 1988, Section 305, which directed the EPA to provide technical assistance to the States in assessing radon concentrations in homes. Through this program, the EPA provided assistance to states in the selection and testing of a probability-based sample of houses. Research Triangle Institute (RTI) supported EPA and the states in this effort during the six years of surveys. Assistance was provided in survey design, interviewer training, sample selection, data processing, and data analysis. In addition, the Agency provided the charcoal canisters used in the surveys and also provided all laboratory analysis.

The goals of the state radon surveys were twofold. Some measure of the distribution of radon levels among residences was desired for major geographic areas within each state and for each state as a whole. In addition, it was desired that each state survey would be able to identify areas of potentially high residential radon concentrations ("hot spots") in the state, enabling the state to focus its attention on areas where indoor radon concentrations might pose a greater health threat.

To ensure the discovery of elevated radon concentrations within a home, the charcoal canisters were exposed under closed-house conditions during the winter and were placed on the lowest livable level. Thus, the estimates of indoor radon concentration provided by the surveys reflect a worst-case scenario and maximize the likelihood of identifying residences with high radon concentrations. The screening measurement provides a measurement of the maximum concentration to which occupants may be exposed. A screening measurement also provides a basis for determining whether additional measurements are needed for making a mitigation decision. Data from these state surveys should not, however, be used directly in assessing health risks, because the screening measurements may overstate annual average concentrations in living areas of these homes.

Since the winter of 1986-87, the EPA has assisted 42 states in conducting surveys of indoor ^{222}Rn concentrations. The 42 states and 6 Indian lands radon surveys included in the National Radon Database were carried out during the six years of the program as listed in Table 1-1. Probability-based surveys also were conducted in six selected Indian lands during four of the six years of the program. The use of probabilities in making house selections allows the results to be extrapolated beyond the sample itself to a well-defined population of homes through the use of sampling weights, which are included in the database for all surveys except Colorado and Connecticut.[1] The sampling weights should be used as described in this documentation to replicate the population estimates presented here. In addition, sample data from state surveys conducted by Colorado and Connecticut are included in the Year 1 database. The sampling weights for these states are set to a value of 0 in the database.

A two-day deployment of open-faced charcoal canisters was used by 24 states and 3 Indian lands during the first three years of the state radon survey assistance program. During these years, a diffusion barrier charcoal canister was developed specifically to be less sensitive to the effects of humidity and air flow than the open-faced canister. Two-day deployment of barrier canisters was used by the eight states and two Indian lands in Year 4 of the program. The exposure period for the barrier canisters was increased from two days to seven days for Years 5 and 6. All devices were analyzed promptly at the EPA laboratory in Montgomery, Alabama. Estimates of the relative measurement error as a percentage of the measured concentration were provided by the laboratory and are included in the database. The performance of the charcoal canisters was monitored periodically through the use of unexposed canisters, canisters exposed to known levels of ^{222}Rn, and collocated canisters.

The database now contains data on short-term screening measurements made on the lowest livable level of over 63,000 randomly selected houses during the winter heating season. Survey results for the 42 states and 6 Indian lands are listed in Table 1-2, which

[1] Colorado and Connecticut conducted state surveys and these data are included in the database for Year 1. Because sampling weights could not be determined for these samples, the survey results for these two states should not be extrapolated beyond the sample. The States of Delaware, Florida, New Hampshire, New Jersey, New York and Utah also have conducted their own surveys. Information concerning these state surveys is included in Appendix D.

shows for each state and Indian land the number of homes tested, the estimated number of residences in the target population, population estimates of the arithmetic mean (average) screening measurement radon concentration, and the estimated population percentage of homes with screening measurements over 4 pCi/L and over 20 pCi/L. Due to the lack of sampling weights for Colorado and Connecticut, reported results are applicable only to the sample households. Results are reported separately for the six Indian lands included in the database.

The geographical distribution of estimated mean screening-level radon concentrations is depicted in Figures 1-1 and 1-2 for the 38 states in the contiguous U.S. with probability-based survey results. These states contain 225 sub-state regions. In Figure 1-1 the regions are grouped into three categories using the estimated regional mean screening measurement: 0 to 2 pCi/L; 2 to 4 pCi/L; and greater than 4 pCi/L. In Figure 1-2, the top 60 regions with an estimated mean screening level over 4 pCi/L are displayed in three more-detailed categories: 4 to 6 pCi/L; 6 to 8 pCi/L; and greater than 8 pCi/L.

Figure 1-3 shows a map of the 10 EPA regions used to define the target population for the surveys of Indian lands. The names and addresses of the EPA regional office radon contacts are included in Appendix D.

Table 1-1 Summary of Six Years of the EPA/State Residential Radon Surveys

Year 1, 1986-87 heating season: ten states

Alabama	(AL)	Michigan	(MI)
Colorado	(CO)	Rhode Island	(RI)
Connecticut	(CT)	Tennessee	(TN)
Kansas	(KS)	Wisconsin	(WI)
Kentucky	(KY)	Wyoming	(WY)

Year 2, 1987-88 heating season: seven states and one Indian land

Arizona	(AZ)	Minnesota	(MN)
Indiana	(IN)	Missouri	(MO)
Massachusetts	(MA)	North Dakota	(ND)
Region 5 Indian Land	(R5)	Pennsylvania	(PA)

Year 3, 1988-89 heating season: eight states and two Indian lands

Alaska	(AK)	New Mexico	(NM)
Georgia	(GA)	Ohio	(OH)
Iowa	(IA)	Vermont	(VT)
Maine	(ME)	West Virginia	(WV)
Region 6 Indian Land	(R6)	Region 7 Indian Land	(R7)

Year 4, 1989-90 heating season: nine states and two Indian lands

California	(CA)	Nevada	(NV)
Hawaii	(HI)	North Carolina	(NC)
Idaho	(ID)	Oklahoma	(OK)
Louisiana	(LA)	South Carolina	(SC)
Nebraska	(NE)	Navajo Nation	(RN)
Billings, MT IHS Area	(RB)		

Year 5, 1990-91 heating season: six states and one Indian land

Arkansas	(AR)	Mississippi	(MS)
Illinois	(IL)	Texas	(TX)
Maryland	(MD)	Washington	(WA)
Eastern Cherokee Nation	(RC)		

Year 6, 1991-92 heating season: two states

Montana	(MT)	Virginia	(VA)

Table 1-2 EPA/State Residential Radon Survey Results, Years 1 to 6

State/Indian Land	# Homes Tested	Estimated # Homes in Population	Screening-Level Estimates Arithmetic Mean	Percent > 4 pCi/L	Percent > 20 pCi/L
AK	1,127	38,287	1.7	7.7	0.6
AL	1,180	565,603	1.8	6.4	0.3
AR	1,535	411,395	1.2	5.0	0.3
AZ	1,507	481,861	1.6	6.5	0.1
CA	1,885	2,232,780	1.0	2.4	0.1
CO*	1,443	1,443	5.2	41.5	2.7
CT*	1,451	1,451	2.8	18.5	0.9
GA	1,534	826,452	1.8	7.5	0.0
HI	523	67,044	0.2	0.4	0.0
IA	1,381	593,815	8.9	71.0	7.5
ID	1,266	187,124	3.3	20.3	1.1
IL	1,450	1,537,325	2.9	19.2	0.8
IN	1,914	992,634	3.7	28.5	1.5
KS	2,009	509,496	3.1	22.5	0.7
KY	879	585,655	2.7	17.1	1.5
LA	1,314	432,162	0.5	0.8	0.0
MA	1,659	1,010,301	3.4	22.7	1.3
MD	1,126	761,456	3.1	18.9	1.4
ME	839	236,917	4.1	29.9	1.9
MI	1,989	1,519,962	2.1	11.7	0.4
MN	919	966,496	4.8	45.4	1.4
MO	1,859	998,706	2.6	17.0	0.7
MS	960	352,285	0.9	2.2	0.1
MT	833	151,605	6.0	42.2	4.7
NC	1,290	1,114,747	1.4	6.7	0.3
ND	1,596	194,315	7.0	60.7	4.3
NE	2,027	310,857	5.5	53.5	1.9
NM	1,885	191,090	3.2	21.8	0.8
NV	1,562	93,004	2.0	10.2	0.8
OH	1,734	1,843,743	4.3	29.0	2.8
OK	1,637	538,309	1.1	3.3	0.0
PA	2,389	2,262,234	7.7	40.5	7.9
RI	376	165,646	3.2	20.6	1.9
SC	1,089	505,281	1.1	3.7	0.3
TN	1,773	741,551	2.7	15.8	1.3
TX	2,680	2,216,326	1.0	3.6	0.2
VA	1,156	972,708	2.3	13.9	1.2
VT	710	117,523	2.5	15.9	0.9
WA	1,935	711,965	1.7	8.8	1.3
WI	1,191	933,700	3.4	26.6	0.8
WV	1,006	324,038	2.6	15.7	0.8
WY	777	74,234	3.6	26.2	1.8
SUBTOTAL	59,395	28,773,526			
R5	934	5,328	2.9	19.7	1.3
R6	740	5,443	2.7	16.9	0.8
R7	669	8,478	5.4	34.9	2.7
RB	187	5,834	2.9	22.3	0.0
RC	594	786	0.8	1.7	0.0
RN	772	33,354	1.7	8.3	0.0
SUBTOTAL	3,896	59,223			
TOTAL	63,291				

(*) - Colorado and Connecticut results apply only to those homes tested in the survey.

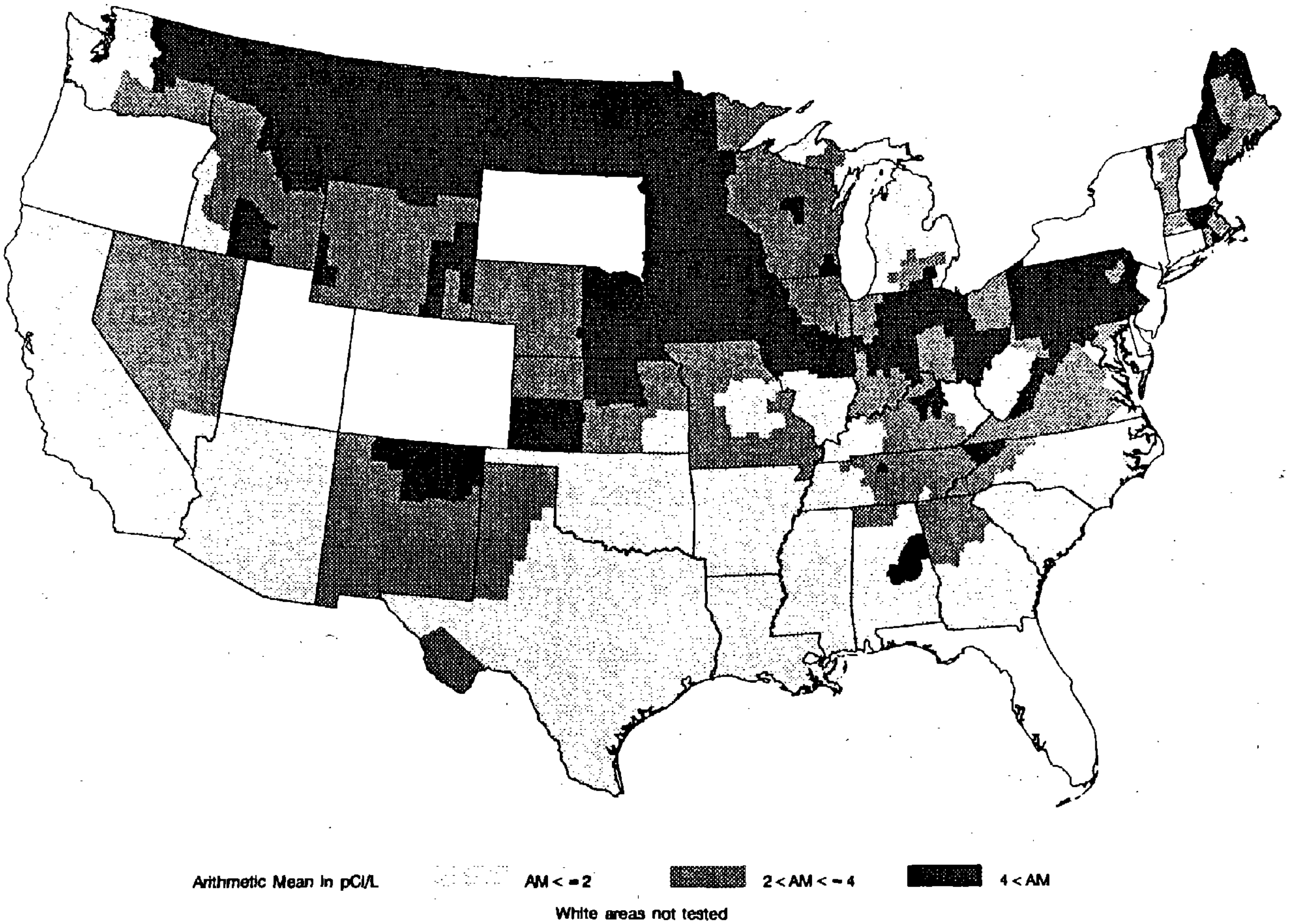


Figure 1. Distribution of Arithmetic Means of Screening Measurements in 225 Regions

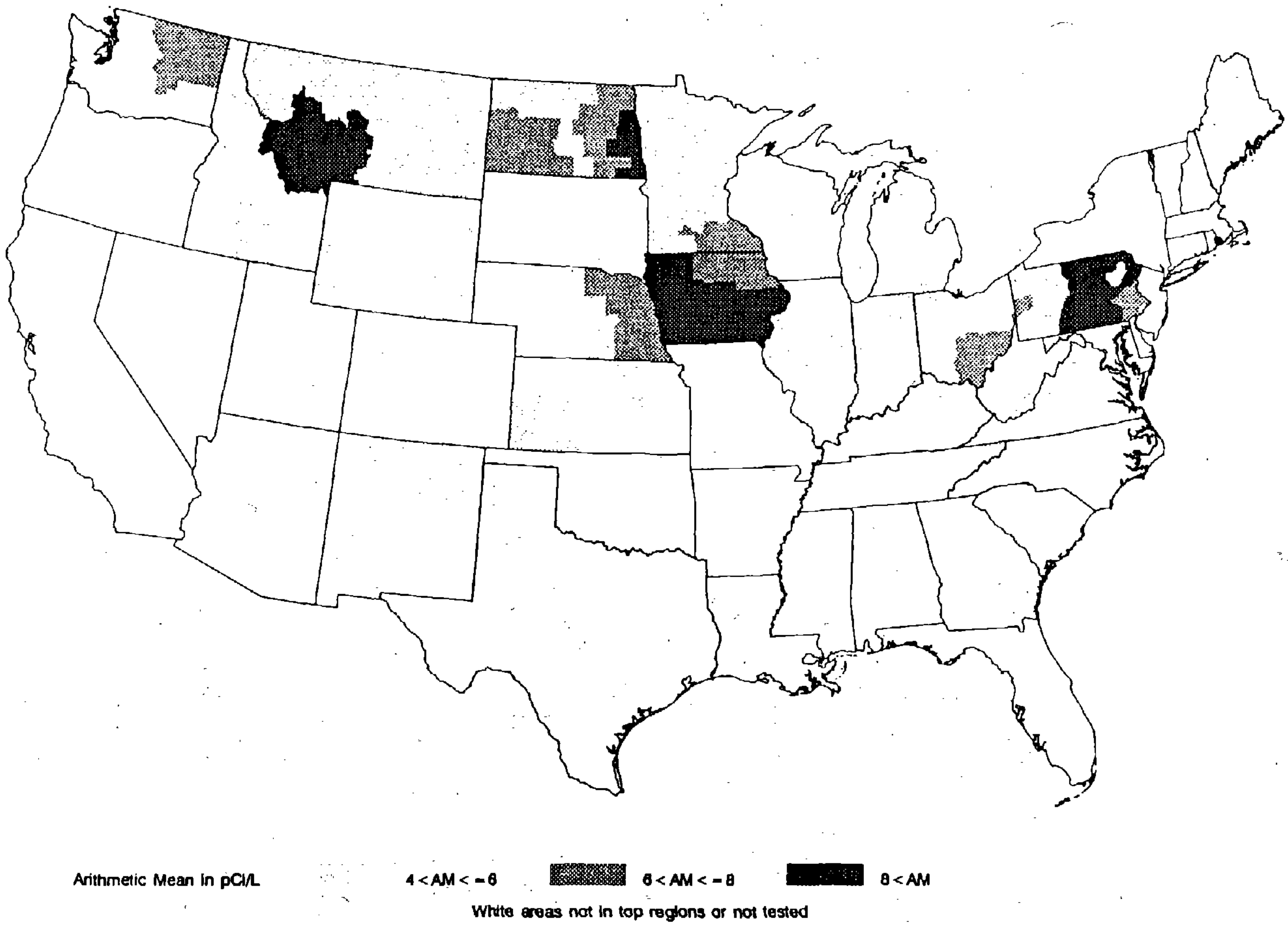


Figure 2. Distribution of Arithmetic Means of Screening Measurements in the Top 60 Regions

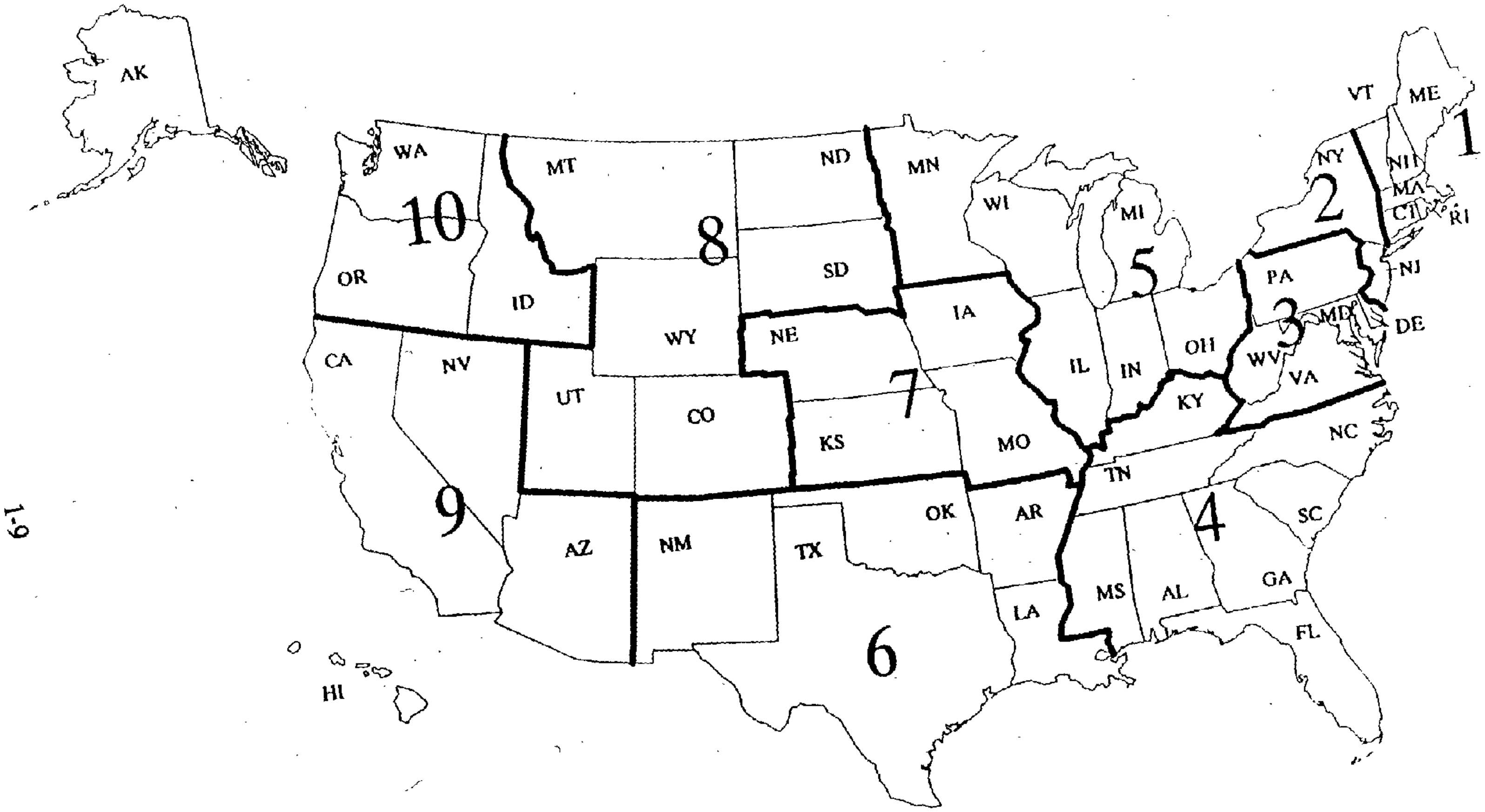


Figure 3. EPA Regions

1.2 SUMMARY OF THE YEAR 3 SURVEYS

During the winter and spring of 1988-89, eight state residential radon surveys and two Indian lands surveys were conducted. The latter two surveys covered Indian lands in eight states located in EPA Regions 6, 7, and 8.

The following 10 state/EPA residential radon surveys were included in Year 3:

Alaska	(AK)	EPA region 6 and 8 Indian lands located in New Mexico, Colorado, Utah, and Texas	(R6)
Georgia	(GA)		
Iowa	(IA)		
Maine	(ME)		
New Mexico	(NM)	EPA region 7 and 8 Indian lands located in North Dakota, South Dakota, Nebraska, and Iowa	(R7)
Ohio	(OH)		
Vermont	(VT)		
West Virginia	(WV)		

For each of the eight states conducting surveys during Year 3, a random sample of residences with listed telephone numbers was selected. For the survey of Indian land, a probability sample of residences was selected for the survey from a listing of all residences located on specified Indian reservations in a total of eight states in Region 6, 7, and 8. Although the sample for Indian lands was selected without regard to the existence of a listed telephone number, information on telephone status was obtained from those selected into the sample.

For each of the Year 3 states, the sample for the state radon survey was a stratified random sample of directory-listed telephone numbers. The first step in designing a survey for a state was to partition the state into three or more geographic regions for which the state wished separate statistical estimates from the survey data. These geologic groupings were then used as strata for sample selection purposes. The states were also asked to identify any areas where the residential radon levels were likely to be high. These areas were samples at a higher rate when necessary to provide good coverage of geographic areas that were suspected of having a radon problem. For convenience in selecting the sample of telephone numbers, county boundaries were used to delineate the geographic regions.

The homes to be measured for indoor radon concentration were selected as follow. First, a probability sample of residential telephone numbers was selected from a sampling frame constructed from the telephone directories for all communities in the state. Telephone numbers in some strata were sampled at higher rates than those in other strata in order to ensure sample sizes large enough to provide precise estimates for each of the designated reporting regions. After the sample was selected, it was partitioned into sample waves, each consisting of a random subsample of 50 telephone numbers. The sequentially numbered waves were implemented in a specified numerical order, permitting the generation of statistical estimates for the random subpart of the sample represented by the implemented waves.

Proceeding sequentially from wave to wave, telephone calls were made to the sample residential telephone numbers. The interviewer first screened for survey eligibility, which required that the dwelling have a floor on or below grade level and, for reasons of liability, that it be owner-occupied. Once survey eligibility was established, the owner-occupant was requested to participate in the survey. Descriptive material about radon and about the survey was provided either before or after solicitation of cooperation. Those agreeing to participate were provided with a canister and instructions for its use, either by mail person. Participants, after exposing the canister for 48 hours, sent it together with a short questionnaire describing where and when the readings had been taken, to the EPA Laboratory in Alabama.

The state radon screening survey results are statistically valid at the state and sub-state regional level. The assignment of counties to regions within each state is detailed in Table C-1 of Appendix C. The number of radon detectors (charcoal canisters) also is shown for each county in this table. Table 1-3 contains population estimates for selected

Table 1-3 Parameter Estimates from the Distribution of Indoor Radon Screening Measurements in Year 3 Surveys, by State and Region (1988-89)

	Number Houses Tested	Est. No. Houses in Population	Arith. Mean pCi/L	Geo. Mean pCi/L	Median pCi/L	75th Percentile pCi/L	90th Percentile pCi/L	% Houses > 4 pCi/L	% Houses > 20 pCi/L
					Alaska				
State	1,127	38,287	1.7	0.7	0.7	1.6	3.3	7.7	0.6
Region 1	282	17,424	1.0	0.6	0.6	1.3	2.1	2.9	0.0
Region 2	325	8,183	3.8	1.8	1.7	2.8	6.6	17.2	2.7
Region 3	267	6,620	0.4	0.2	0.2	0.5	0.8	1.5	0.0
Region 4	253	6,059	2.2	0.9	1.1	2.6	5.3	15.3	0.4
					Georgia				
State	1,534	826,452	1.8	1.2	1.3	2.1	3.6	7.5	0.0
Region 1	872	552,677	2.1	1.5	1.5	2.5	4.1	10.3	0.0
Region 2	423	211,029	1.2	0.9	1.0	1.6	2.3	2.6	0.0
Region 3	239	62,747	0.8	0.5	0.7	1.0	1.3	0.3	0.0
					Iowa				
State	1,381	593,815	8.9	6.1	6.9	11.4	17.4	71.0	7.5
Region 1	144	44,385	9.8	7.9	7.9	11.3	15.9	86.9	4.2
Region 2	160	49,236	10.2	7.7	8.5	12.8	18.0	82.1	8.2
Region 3	128	39,432	9.1	6.8	7.9	12.8	16.6	71.7	5.5
Region 4	138	42,626	7.4	5.4	5.4	9.4	16.2	67.5	5.1
Region 5	187	85,341	7.7	5.0	5.4	8.1	13.9	62.5	4.9
Region 6	179	129,810	10.3	7.6	7.9	13.3	19.2	81.0	9.2
Region 7	169	114,649	8.1	5.1	5.9	10.4	15.2	61.4	9.0
Region 8	143	32,209	8.7	5.5	7.1	10.9	15.1	67.6	6.3
Region 9	133	56,126	8.8	5.2	5.4	9.8	18.9	61.7	10.3
					Maine				
State	839	236,917	4.1	2.2	2.2	4.7	9.1	29.9	1.9
Region 1	211	73,190	5.6	3.2	3.1	6.6	14.0	39.9	3.4
Region 2	108	37,430	3.3	2.1	2.0	3.8	7.7	25.7	0.0
Region 3	95	27,248	4.7	2.2	1.9	4.6	9.4	33.7	3.2
Region 4	109	31,293	2.6	1.6	1.7	2.9	6.9	19.1	0.0
Region 5	93	17,090	3.0	1.7	1.8	3.7	7.3	22.6	0.0
Region 6	121	31,861	2.3	1.4	1.7	2.9	5.1	16.6	0.3
Region 7	102	18,806	4.8	3.0	3.5	6.3	10.5	40.9	4.8
					New Mexico				
State	1,885	191,090	3.2	2.2	2.2	3.6	6.1	21.8	0.8
Region 1	855	98,158	3.4	2.5	2.4	3.7	6.1	23.5	0.9
Region 2	422	24,041	4.7	3.2	3.4	5.7	9.5	41.6	2.0
Region 3	316	29,466	2.2	1.5	1.4	2.5	4.2	11.2	0.7
Region 4	292	39,425	2.3	1.6	1.7	3.0	4.9	13.3	0.0

Table 1-3 Parameter Estimates from the Distribution of Indoor Radon Screening Measurements in Year 3 Surveys, by State and Region (1988-89) (Continued)

	Number Houses Tested	Est. No. Houses in Population	Arith. Mean pCi/L	Geo. Mean pCi/L	Median pCi/L	75th Percentile pCi/L	90th Percentile pCi/L	% Houses > 4 pCi/L	% Houses > 20 pCi/L
					Ohio				
State	1,734	1,843,743	4.3	2.2	2.1	4.8	9.2	29.0	2.8
Region 1	445	317,473	4.7	2.7	2.8	5.7	10.0	36.7	2.7
Region 2	475	726,722	3.2	1.6	1.6	3.1	6.7	18.3	2.2
Region 3	386	485,903	4.0	2.1	2.0	4.4	8.5	27.3	2.1
Region 4	428	313,647	7.0	3.5	3.8	7.6	14.0	48.4	5.6
					Vermont				
State	710	117,523	2.5	1.3	1.2	2.6	5.9	15.9	0.9
Region 1	164	40,290	2.0	0.9	0.8	1.6	4.1	10.8	1.3
Region 2	130	16,293	2.6	1.6	1.7	2.9	5.3	16.7	0.7
Region 3	130	20,805	2.9	1.6	1.7	3.3	6.9	19.5	0.9
Region 4	128	20,293	2.3	1.2	1.1	3.0	6.0	19.8	0.0
Region 5	158	19,841	3.2	1.7	1.7	3.5	6.5	18.0	1.2
					West Virginia				
State	1,006	324,038	2.6	1.4	1.4	2.8	5.4	15.7	0.8
Region 1	211	57,837	4.7	2.2	2.1	4.8	9.1	31.2	3.4
Region 2	240	62,761	3.0	1.7	1.8	3.2	6.2	17.6	1.2
Region 3	555	203,440	1.9	1.2	1.1	2.2	4.1	10.7	0.0
					Region 6 Indian Lands				
All	740	5,443	2.7	1.7	1.8	3.1	5.3	16.9	0.8
Region 1	236	2,090	2.6	1.4	1.5	2.9	5.4	14.9	1.6
Region 2	123	903	2.4	1.8	1.8	2.7	4.1	13.0	0.0
Region 3	150	488	3.0	1.5	1.6	2.7	6.3	16.5	2.4
Region 4	111	984	2.4	2.0	1.9	2.7	4.2	12.8	0.0
Region 5	120	978	3.5	2.7	2.5	4.2	6.5	28.9	0.0
					Region 7 Indian Lands				
All	669	8,478	5.4	2.8	2.9	5.4	10.3	34.9	2.7
Region 1	180	2,342	5.6	3.6	3.6	6.7	13.0	44.2	2.3
Region 2	148	1,944	4.3	2.5	2.7	5.4	7.8	31.9	1.5
Region 3	202	2,513	4.5	1.9	1.9	3.6	5.6	19.5	1.5
Region 4	75	968	10.2	4.6	4.9	10.8	19.1	56.5	11.1
Region 5	64	711	4.1	3.0	3.3	4.7	7.6	37.3	0.0

parameters of the regional and state-wide radon distribution. These estimates were obtained using the appropriate sampling weights, as described in Section 3.3. The table contains estimates of the mean (average) screening measurement, the median, the geometric mean, the 75th and 90th percentiles, and the percent of houses over 4 pCi/L and over 20 pCi/L.

In summary, each state radon survey is designed to provide statistical estimates of radon concentration

- In owner-occupied residences,
- With listed telephones numbers, and
- A floor at or below ground level.

Each survey of Indian lands is designed to provide estimates of radon concentrations in owner-occupied residences having a floor at or below ground level.

2. The Sample Design

2.1 THE OVERALL SAMPLING PLAN

The sampling plan for the state radon surveys called for the selection of probability samples of residences in each state. A probability sample is one in which every element in the population has a positive chance of selection, and, for every element in the sample, the selection probability or relative probability is known. Probability sampling permits the extrapolation of survey results to the entire population and, in addition, can permit the calculation of measures of precision for the estimates. Because one of the goals of each state radon survey was the generation of estimates of distributions of residential radon levels for eligible residences in the state as a whole and for the major geographic areas within the state, use of probability sampling was imperative. Probability-based surveys were also necessary to validly compare results from one state with results from another.

2.2 POPULATION DEFINITION AND SAMPLING FRAMES

The target population for the surveys in all eight of the Year 3 states consisted of owner-occupied residences with a permanent foundation and at least one floor at or below ground level and with a telephone number published in the latest directory. (Mobile homes with permanent foundations on airtight panels/skirts and with a published telephone number are also included.) The statistical estimates generated from the survey data apply to this population.

In reality, the totality of occupied residences in the state constituted the population of interest. However, as is often the case in survey research, surveying this population was not deemed feasible, for several reasons. First, it was considered inadvisable from a legal point of view to include rental dwellings without first obtaining the permission of the owner. Although procedures could be devised to obtain such permission, the cost of doing so, both in dollars and in delay in the survey schedule, was deemed impractical. Second, homes that had no floor on or below ground level were excluded from the survey target population. Although these homes are no doubt usually rental apartment units, the category would include some owner-occupied condominiums. These were excluded from the target population because radon levels on upper floors were expected to be low, and it was felt that the focus of the survey should be on residences that were potentially at risk. Third, the survey target population was restricted to homes with listed telephone numbers, basically because of time and cost considerations. Sampling of homes without regard to the existence of a telephone would call for an area probability procedure, which requires onsite staff for both listing and data collection and is both expensive and time consuming. The telephone survey approach was used because it offered a more economically feasible alternative. Telephone surveys can be implemented using a relatively small staff working in a central location, and they can be carried out on short notice and within a restricted time schedule.

Two types of samples are commonly used for telephone surveys: random digit dialing samples, for which every possible telephone number is given a positive chance of being selected into the sample, and telephone directory samples, for which only listed telephone numbers are given a chance of selection. In both Year 1 and Year 2, each state was given the choice of these two telephone survey methods, and each chose the procedure calling for the selection of listed telephone numbers. There were two major incentives for making this choice. First, the labor involved in telephoning is much less using listed telephone numbers than it is using random digit dialing because the vast majority of listed numbers will be working residential numbers, as compared to only about 20 percent for the random digit dialing technique. Second, names and addresses are available for directory-selected telephone numbers, making possible a mailing of material describing the health risks associated with radon exposure and describing the survey. This second reason was an important consideration for those states wishing to do a mailing prior to the telephone contact.

The survey procedures, forms, training manuals, ID numbering schemes, and data entry program were developed for Year 1 survey and refined for Year 2 surveys, based on samples selected from directory files. Because of this heavy investment, coupled with the apparent state preference for directory samples, the Year 3 states were not offered the option of a random-digit-dialing methodology. All Year 3 state radon surveys used samples selected from directory-based files.

Two organizations constructed files of listed telephone numbers: Survey Sampling and Donnelley Marketing. While both organizations had comparable sampling frames, Survey Sampling was more restrictive in the selection procedures that they were willing to implement. Only Donnelley Marketing was willing to follow precisely the sample selection procedures developed for the state radon surveys; therefore, samples for all eight Year 3 states were purchased from that company.

2.3 STRATIFICATION AND SAMPLE ALLOCATION

To improve the precision of the survey estimates, the sampling frame for each of the six states using the Donnelley frame was stratified prior to sample selection. Because different sampling rates can be used for different strata, it was possible to control the size of the sample to be selected from each reporting region. Two or more alternative sampling allocations were produced and provided to each state. The first allocation was based on equal probability sampling, which yields samples that are distributed across strata in the same way the population is distributed. One of the disadvantages of equal probability sampling is that it can result in small sample sizes for small reporting groups.

The second alternative allocation that was provided avoided this potential problem by allocating the sample equally to the different strata. However, to achieve an equal allocation when the strata vary in size, different sampling rates must be used for the different strata. The unequal sampling weights, which must be used in the estimation process in order to account for the differing sampling rates, can have the effect of lowering the precision of the statewide estimates.

There are obvious tradeoffs among the different allocation alternatives. For each allocation provided to a state, a table showing the expected precision for statewide and reporting group estimates was provided. This enabled the state to view the tradeoffs in precision associated with the different types of allocations.

States are usually interested in the expected distribution of the sample across the counties in the state. For each of the sample allocations, a distribution showing the expected sample size for each county was therefore produced using the Market Statistics' estimate of the number of occupied housing units in each county in 1989. The expected sample take was obtained by applying the sampling rate for the stratum to the estimated number of occupied housing units.

Prior to the fall 1988 orientation meeting, each Year 3 state was provided with descriptive information about the proportional allocation, based on equal probability sampling, and the equal allocation. The descriptive information consisted of precision and distribution tables, described above, as well as a discussion of the advantages and disadvantages of each allocation. The state representatives were therefore able to consider two sample designs prior to participating in the detailed survey planning sessions that were carried out for each state survey at the conclusion of the orientation meeting. Tables for additional allocations were prepared when appropriate so that the state could see the effect of increasing or decreasing the overall size of the sample, the effect of sampling more heavily in sparsely settled areas, or the effect of sampling more heavily in areas that were suspected of having elevated residential radon levels.

After considering all of the location options provided, the state, with EPA's approval, decided on one of the allocations.

A description of the allocation that was chosen by the state, the target number of canisters to be placed, the sampling rates used in the strata, and the expected design effect (DEFF) due to unequal weighting for variables that are uniformly distributed across strata are presented for each state in Appendix C.

Following guidelines determined by the agreed upon allocation, the samples for the eight states were selected from the Donnelley Marketing files. In all cases, detailed instructions for ordering the file and selecting the sample for each state were prepared. The instructions called for ordering the residential telephone listings in each stratum by the size rank of the county in which the residence was located, then by the census block group or enumeration district. The listings were finally ordered by telephone number. This ensured maximum geographic spread when systematic random sample selection procedures were used.

2.4 SAMPLE SELECTION PROCEDURES FOR THE DONNELLEY FILE SELECTIONS

To permit the unbiased estimation of the sampling errors of the survey estimates of radon characteristics for the state and for major geographic subparts of the state, five independent, systematic random samples were selected from each stratum. To do this, RTI provided the sample size to be selected from each stratum for each of the five samples, a list of the counties that made up each stratum, and the specifications for ordering the file within each stratum. The sample selection instructions that were provided by RTI are presented in Table 2-1.

The following variables were requested for each sample selection:

1. State code from the Federal Information Processing Standards (FIPs),
2. County FIPS code,
3. Stratum,
4. Area code,
5. Telephone number,
6. Name,
7. Mailing address,
8. ZIP code, and
9. Sample (or replicate) number (1-5).

Table 2-1 Procedures for Selecting the Sample of Telephone Numbers

1. Sort all residential telephone numbers in the state as specified.

2. Determine the number of listings of residential telephone numbers on the file for the stratum Call this number L.

3. Identify the sample size specified for the stratum and call this number S.

4. Divide L by S and round to the nearest whole number to obtain the Selection Interval I.

5. Select five different random numbers between (and including) 1 and I.

6. Successively add I to the first random number to generate approximately S selection numbers for the stratum to identify the sample telephone numbers of the ordered list.

7. Repeat step 6 for each of the other four random numbers until all five random samples of size S have been selected.

8. When this procedure has been implemented for all strata defined for a state, the state's sample selection is completed.

2.5 THE ALASKA SURVEY

The Alaska survey was handled somewhat differently from the other state radon surveys because that state had special problems related to the accessibility and the construction characteristics of the homes.

First, most of the population of the state is concentrated in and around the three largest cities--Anchorage, Fairbanks, and Juneau. However, the remaining population is spread over a vast geographic area, most of which is not easily accessible. In many of the remote villages, few homes have telephones, and mail service during the winter months can be very undependable. Even though these areas could obviously not be covered by a telephone survey, the state, nevertheless, wished to have some information about the residential radon level there.

A second characteristic about Alaska that affected the survey was the type of construction of the homes in the northern part of the state. Because these homes were built above the ground, with no permanent skirting, they are not subject to a potential elevated radon problem, and the state did not wish to include them in any survey. In addition some areas in the state were military installations, which also were to be excluded from any survey.

The Alaska residential radon survey was designed to handle the circumstances described above. Certain areas were excluded from the survey chiefly because they contained only unskirted homes above ground level or because they were "mostly military." The telephone exchanges in the areas of the state that were to be surveyed were classified either as "telephone survey," for those exchanges serving the area in and around the three largest cities, or as "personal visit," for the remaining more remote areas.

The telephone survey component of the study was designed to yield statistical estimates for the defined target population in and around the three largest cities. The same methods were used for the Alaska telephone survey as were used in other state surveys. The listed telephone numbers in the telephone exchanges to be covered in the telephone survey were partitioned into strata and ordered, as described earlier in this chapter. Five independent probability samples of residential telephone numbers were then selected. The sample numbers were called and screened for survey eligibility, which followed the same definition as that used in other states. Charcoal canisters were mailed by participants to the EPA Laboratory in Montgomery, Alabama.

The personal visit survey Component of the study was designed to yield statistical estimates for the totality of purposively selected "remote villages." The methods used for this component were specially designed to fit the unusual situation in Alaska. A village was to be included in the personal visit survey if and when a government employee visited the village on government business (not related to the survey). Upon arrival in the village, the government employee was to prepare a listing of all of the homes in the village, assign random numbers that in effect placed the listings in a random order, and select as many of the randomly ordered homes as needed to find five that were survey eligible. (Survey eligibility for this component of the survey required the home to be owner occupied and have a floor on or below ground level. There was no requirement to have a telephone.) Charcoal canisters would be placed in the five randomly selected homes and would be retrieved 48 hours later. Upon return, the government employee was to mail the retrieved exposed detectors to the EPA Laboratory in Alabama for reading.

Because of the great travel expense involved, villages would be included in the survey only if a government employee was visiting the village for some other purpose. A sample size of five was set because it was felt that five homes was the maximum number that could be included without unduely extending any visit to the village. This personal survey component would not support statistical estimates for all remote villages because the villages themselves would not have been randomly selected. However, it was expected that the personal survey component would support statistical estimates for the totality of villages that were surveyed.

Unfortunately, the unusually cold weather in Alaska during the 1988-89 winter delayed travel and delayed mail service. As a result of the delay in receiving the exposed canisters, the EPA Laboratory was unable to assign readings to many of the canisters they received from the remote villages. Therefore, no statistical estimates were produced from this component of the survey, and all data from this source must necessarily be treated as being comparable to the data obtained from volunteers.

2.6 PARTITIONING THE SAMPLES INTO WAVES

Estimating the exact number of sample selections that would be needed in a state survey to be able to place the desired number of canisters was very difficult. Unknown were the exact proportion of selected numbers that would be working residential numbers, the exact proportion of residential numbers that would be associated with survey-eligible residences, or the proportion of eligible residences that would participate in the study. Another very important unknown was when the weather in the state would become so warm that the closed house requirement for canister deployment could not be met, and the survey would have to be discontinued.

There is a commonly used technique for controlling the number of survey participants in situations where many unknowns are involved in estimating the number of sample selections needed. The procedure involves partitioning the sample into a number of random subsamples and implementing only as many of the subsamples as are needed to achieve the desired number of participants. This technique was used in all eight Year 3 states.

A sample sufficiently large for any reasonable set of assumptions was selected as described above. It was then partitioned into random subsamples, or waves, of 50 telephone listings each. The waves were randomly ordered and numbered sequentially, and they were activated in a specified numerical order by the states. Implementation of the sample in random subparts meant that a state did not need to complete all sample waves.

The procedures used in processing the file and partitioning the sample into waves are as follows:

1. The sample of 10-digit telephone numbers was checked for duplicates, which were eliminated, and was checked to verify that the proper number of records had been provided for each replicate in each stratum.

2. The total number of waves, W, into which the sample was to be partitioned was determined by dividing the number of records on the file by 50.

3. The wave numbers 1 through W were put in random order and assigned to the first W records on the file, which had been ordered by stratum, replicate, and telephone number. The wave numbers 1 through W were again placed in a random order and assigned to the second W records on the ordered file, etc., until each record had been assigned a wave number.

4. The records were ordered by wave number, and a Case ID number was assigned sequentially.

2.7 SAMPLE SELECTION PROCEDURES FOR THE INDIAN LANDS SURVEY

The Region 6 IHS carried out a residential radon survey on Indian lands in Colorado, Utah, New Mexico, and Texas and the Region 7 IHS carried out a similar survey on Indian lands in North Dakota, South Dakota, Nebraska, and Iowa. For both surveys, a personal interview procedure was used rather than a telephone interview procedure, as had been used for the other surveys. The canisters were also placed and retrieved by the field interviewer.

The target populations were also somewhat different for the Indian lands survey. All owner-occupied homes, with at least one floor on or below ground level and located on one of the Indian lands survey locations were eligible for the survey, whether or not the land on which the house was located was owned by the occupant and whether or not there was a listed telephone number linked to the home. However, each respondent was asked if there was a telephone at the home and if the telephone number was listed. This information permitted partitioning the Indian lands sample into two groups: those cases that could be similar in definition to cases in the state sample and those that did not comply with the eligibility definition used in the state sample. This partitioning would permit a state to produce consistent state-level estimates that include homes on Indian lands.

The IHS constructed a sampling frame for each reservation, noting the name and address of each family living on the reservation. For ease in distributing canisters to each of the reservations and for controlling the overall sample size, each reservation, denoted by the subscript h, was assigned a specific sample size, n'_h. The n'_h constituted the expected sample sizes that would provide the desired distribution and total number of sample cases across reservations. The sequential numbers representing addresses within each reservation were put in a random order, and the first n'_h addresses on the list were assigned to the primary sample. The following 1/2 n'_h addresses on the list were assigned to the secondary sample, which was to be used, in the order specified, as needed. The IHS staff used the selected sequential numbers to identify the addresses of the sample selections.

The interviewer visited all of the n'_h cases in the primary sample, determined survey eligibility, and attempted to place a canister in each eligible home. Some primary sample cases were found to be ineligible for the survey. If, for example, the family had moved from the reservation and left their reservation home vacant, the sample case would be classified as "not survey eligible." On some occasions, a refusal was obtained for survey-eligible home in the primary sample. Whenever participation was obtained from fewer than n'_h eligible homes in the primary sample, the secondary sample cases were worked in the order assigned until detectors were placed in exactly n'_h eligible homes on the reservation.

3. Estimation Using Survey Results

3.1 CALCULATION OF SAMPLING WEIGHTS

Because most of the states used unequal probability sample designs for their state radon surveys, sampling weights that account for the unequal probabilities of selection must be used to generate unbiased population estimates from the survey data. Sampling weights that reflect only the differential selection probabilities would be adequate if 100 percent response rates and participation rates were achieved. However, this level of response was not obtained. For the state radon surveys, some of the sample cases failed to complete a screening interview, either because they were never successfully contacted or because they refused to provide the screening information. Whether or not they were in fact eligible was, therefore, never determined. For other cases, the screening information was provided, and the housing unit was determined to be eligible for the survey, but a canister reading was not successfully linked to the case. There are numerous reasons why this might have occurred. The canister may not have been read because it was never deployed; it may have been deployed but never returned; or it may have been returned but not received in time to be included in the analysis. In addition, clerical or keying errors associated with matching criteria could have prevented matching canister readings with the proper cases. To compensate for the missing information, a weighting class adjustment was used. This procedure increased the sampling weights of participants to compensate for the missing information from nonparticipants. The steps used in calculating sampling weights and adjustments for the eight Year 3 states are described below.

The first step in calculating the sampling weight was determined from the information provided by Donnelley Market Services. For each stratum in the sample, RTI was provided with the number of listings from which the sample was selected. The number of selections that should be made was specified. Using this information, the first component of the sampling weight was computed for each stratum and used for all selections from that stratum. For any stratum h, the first sampling weight component was calculated as

$$W'_h = N_h / [(5)(n_h)], \tag{1}$$

because five samples of size n_h were selected from the N_h listings in stratum h.

As was described in Chapter 2, each state's sample was randomly partitioned into waves of 50 listings each, each wave being in effect a probability sample of the entire state. Although all waves were available for use in the state radon survey, not all were used. The second component of the sampling weight represented the portion of the sample waves that were included in the analysis. Any wave for which at least 45 of the 50 cases were completed was considered to have been implemented, and it was referred to as an "active" wave. Computer runs were made on the Control/Screening Form file to determine which waves would be classified as "active" and included in the analysis and which would not. For each state, the sampling weight component reflecting the proportion of waves classified as active was computed. This was merely the total number of waves of 50 listings divided by the number of waves classified as active waves, or V/v. Only cases in the v active waves were used in the remaining calculations and in the analysis.

Next an unadjusted sampling weight was calculated for every selected case in every active wave, regardless of the response or participation status of the case. This weight was merely the product of the two weight components.

$$W''_h = (W'_h)(V/v) \tag{2}$$

Next, every record in every active wave was compared to the file of canister readings and, by matching on House ID number, was classified as a participant or a nonparticipant. All active wave cases classed as participant would be used in the analysis, because they were in an active wave and had a canister reading. To adjust for missing canister readings for the survey eligible, all active wave nonparticipant cases were further classified according to eligibility status. The following groups were formed for the active wave cases:

Group A:	Participants (all eligible cases for which a canister reading was available).
Group B:	Survey eligible nonparticipants.
Group C:	Nonparticipants, survey eligibility unknown. (All cases for which eligibility information was never obtained.)
Group D:	Nonparticipants known to be ineligible for the survey.

These four groupings were used in calculating the adjustments for nonresponse.

Five weighting classes were formed within each stratum, corresponding to the five replicates used in the sample selection. Within each weighting class, an adjustment-for-nonresponse factor was computed, as follows.

First, an estimate of the proportion of cases that were survey-eligible was computed:

$$A'_{sh} = \frac{\left|\Sigma W''_{shi}\right|_A + \left|\Sigma W''_{shi}\right|_B}{\left|\Sigma W''_{shi}\right|_A + \left|\Sigma W''_{shi}\right|_B + \left|\Sigma W''_{shi}\right|_D} \qquad (3)$$

where

$\left|\Sigma W''_{shi}\right|_A$ =	sum of the unadjusted sampling weights over all nonparticipants in the s replicate in stratum h,
B =	survey-eligible nonparticipants, and
D =	nonparticipants known to be ineligible.

The proportion A'_{sh} was used to estimate the proportion eligible among those for whom eligibility has not been determined. This figure was needed to determine the nonresponse adjustment factor for each replicate s within each stratum h:

$$A_{shi} = \frac{|\Sigma\, W''_{shi}\, |_A + |\Sigma\, W''_{shi}\, |_B + A'_{sh}\, |\Sigma\, W''_{shi}\, |_C}{|\Sigma\, W''_{shi}\, |_A} \tag{4}$$

where

$|\Sigma\, W''_{shi}\, |_C$ = sum of the unadjusted weights over all nonparticipants with unknown eligibility and where all other terms are as defined above.

The final sampling weight was then calculated for each sample case in every active wave as:

$$W_{shi} = (W''_{shi})(A_{shi}) , \tag{5}$$

and the sampling weight W_{shi} was used as the sampling weight in all analysis.

3.2 CALCULATING SAMPLING WEIGHTS FOR THE INDIAN LANDS SURVEY

A modification of the above procedures was used for the Indian lands survey. A negative binomial distribution was assumed in which n''_h sample homes were contacted on a reservation to obtain n'_h survey-eligible homes. Some of the n''_h selections came from the primary sample, but some could have come from the secondary sample. The proportion of survey-eligible homes for the reservation was estimated to be:

$$(n'_h-1)/(n''_h-1) \tag{6}$$

The number of survey-eligible homes on the reservation was estimated to be:

$$N_h = [(n'_h - 1)/(n''_h - 1)]N'_h \tag{7}$$

where

n'_h = desired sample size from reservation h,

n''_h = number of case that needed to be contacted to discover n'_h survey-eligible residences in reservation h, and

N'_h = number of listings on the reservation h sampling frame.

The final sampling weight was calculated for each of the participants providing a usable detector reading as

$$W_{hi} = N_h/(\#\text{usable participants})_h.$$

3.3 ESTIMATING MEANS AND PROPORTIONS

The analytical results were obtained using SESUDAAN, a computer software program developed by RTI for analyzing survey data with complex error structures. Formulas for estimating means and proportions from the state surveys using this program are shown below. Appendix E contains the formulas for estimating means and proportions for the Indian lands.

Define Y^*_r as the true mean radon level for the r^{th} region or reporting group ($r = 1,...,R$). Y^*_r can be estimated as

$$Y^*_r = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi} Y_{hi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi}} \qquad (8)$$

where

Y_{hi} = observed radon measurement for the i^{th} eligible household in stratum h ($i = 1, ..., n_h$, $h = 1, ..., H$);

W_{hi} = sampling weight associated with Y_{hi}; and

$$J_{rhi} = \begin{cases} 1 & \text{if } i^{th} \text{ eligible household in stratum h is in the } r^{th} \text{ region} \\ 0 & \text{otherwise.} \end{cases}$$

The estimated mean for all regions combined (i.e., the statewide estimate) is given by

$$Y^*_o = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi} Y_{hi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi}} \qquad (9)$$

Similarly, define P^*_r as the true proportion of eligible households in the r^{th} region with radon levels exceeding X pCi/l. P^*_r can be estimated as

$$P^*_r = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi} I_{xhi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} J_{rhi} W_{hi}} \qquad (10)$$

where

W_{hi} and J_{rhi} are as previously defined and

$$I_{xhi} = \begin{cases} 1 & \text{if measurement on } i^{th} \text{ eligible household in stratum h is greater than X pCi/l} \\ 0 & \text{otherwise.} \end{cases}$$

The estimated proportion for all regions combined (i .e., the statewide estimate) is given by

$$P^*_o = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi} \, I_{xhi}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} W_{hi}} . \tag{11}$$

4. Methodological Results

The survey methodology used for the third year if the state/EPA radon survey program was reviewed at five different levels:

- First, the coverage of each state survey was assessed. To do this, four different estimates were compared of the number of owner-occupied single family housing units having a telephone, which was the approximate definition of the survey-eligible population. For each state, the survey estimate of this population size was compared to an estimate based on the 1980 Census counts for the state, to an estimate made using current counts from the Donnelley Marketing Service files from which most of the state samples were selected, and to an estimate based on the Market Statistics' projections.

- Second, the response rate and the participation rate obtained in each of the states were computed. These were simply the ratio of the estimated number of respondents to the estimated number of eligible and the ratio of the estimated number of usable canister readings to the estimated number of eligible.

- Third, the number of cases for which eligibility status was never determined was reviewed.

- Fourth, the Control/Screening Forms that were returned by the states to identify the types of errors that the states made in carrying out the survey were reviewed.

- Fifth, all of the problems that occurred throughout the course of all of the Year 2 state radon surveys were assessed to determine whether modifications were needed in survey procedures.

In the sections that follow, each of these assessments of the state radon survey methodology is discussed.

4.1 COVERAGE

The results of the coverage investigation are presented in Table 4-1. For each of the eight Year 3 states, the number of owner-occupied single family housing units with a telephone was estimated using 1980 decennial census information, using Donnelley file counts, using the Market Statistics' estimates, and using state radon survey results. In constructing these estimates, the percentage of housing units that were owner occupied was available by state, but the percentage of owner-occupied housing units that were single unit structures was available only for the nation as a whole. The national average, showing 94 percent of all owner occupied housing as being single unit structures, was therefore used in the calculations for each of the states. In addition, the nationwide estimate of 97 percent was used for the percentage of owner-occupied single structure housing units having a telephone.

Column 3 of Table 4-1 shows an estimate of the approximate number of survey-eligible housing units using 1980 Census counts, and columns 5 and 9 show comparable estimates made from the Donnelley file counts and Market Statistics' estimates, respectively. The ration of the Donnelley estimates to the Census estimates, shown in column 6, vary from a low of 0.55 for Alaska and 0.71 for West Virginia to a high of 1.06 for Vermont. Column 7 shows comparable ratios for estimates of survey eligible based on Donnelley file counts to those using Market Statistics' data. These ratios vary from a low of 0.37 for Alaska and 0.59 for New Mexico to a high of 0.93 for Vermont. The two sets of ratios were calculated to get a very rough indicator of what might be missing by using the Donnelley files as sampling frames, without using a supplementary procedure for picking up housing units not linked to a Donnelley listing, but otherwise survey-eligible. The relatively low ratios for New Mexico and West Virginia indicate a potential for a sizable noncoverage. The very low ratio for Alaska reflects a lower than national average percentage of homes having a telephone and may also reflect poorer coverage by the Donnelley frame and poorer estimates used in these comparisons. The partitioning of

Alaska into the telephone survey component, the personal survey component, and the excluded component complicates the estimation of population sizes.

Column 15 shows the ratio of the number of survey eligible in each state, as estimated from the survey itself, to the estimate made directly from the Donnelley frame counts. This ratio was calculated as a measure of the loss suffered because of movers and possibly because of households being difficult to reach by telephone. Recall that the procedures selected a sample of telephone numbers and the housing units linked to those numbers, regardless of whether the address was the same as was given in the sampling frame. Therefore, housing units of movers were picked up, but not to the degree in which they were lost. When someone moves, their telephone number is typically retired for a period of 6 months to a year, unless it is carried to the new home. Therefore, a good many movers were reached at their new home. Intrastate movers who change telephone numbers and those who move in from another state were lost if the move occurs after the cutoff date for directories on which the Donnelley listing are based. The ratio of survey-estimated survey eligible to Donnelley-estimated survey eligible ranges from a low of 0.90 for Georgia to a high of 1.00 for Alaska, Iowa, and Ohio, indicating very little loss because of movers.

4.2 RESPONSE RATES

Approximate observed response and participation rates are presented in the bottom two rows of Table 4-2. The percentage of known survey-eligible housing units for which the respondent agreed to place a charcoal canister in the home ranges from a low of 82 percent for Region 6 Indian lands to a high of 97 percent for the Alaska telephone survey and response rates over 90 percent for six other surveys.

Participation rates show the percentage of known survey-eligible homes for which a usable canister reading was obtained. These percentage vary from a low of 66 percent for Ohio to a high of 91 percent for the Region 7 Indian lands survey. The high figure for the latter group represents the success of the personal placement and retrieval procedures used in this survey. The highest participation rate for a state survey was 84 percent for the Alaska telephone survey.

Although the average response rate for known eligible for the 10 Year 3 surveys was about 91 percent, the average participation rate was only about 78 percent, a drop of about 14 percentage points. Getting people to return their canisters immediately after exposing them for the designated period might be an aspect of data collection that will continue to be a given a great deal emphasis. States will continue to be encouraged to recontact people to whom a canister has been sent, but no reading received, to remind them to deploy their canister and to return it immediately after exposure.

4.3 UNKNOWN ELIGIBILITY STATUS

Most of the Year 3 states did an excellent job in returning all Control/ Screening Form for all of their activated waves. This aspect of the data collection process received more emphasis in the Year 2 and Year 3 training because it had been found to be a major problem in Year 1. There does, however, seem to be a large number of "eligibility unknown" cases for several states, and these were especially high for Ohio and Vermont, where about one-fourth of all activated sample cases were so classified. The "eligibility unknown" classification was assigned not only to cases in activated waves for which no screening form was received, but also to cases with repeated ring-no-answer calls and to cases for which a contact was made but the screening interview was not completed to the point where eligibility for the survey could be established. It is extremely important to call on different days of the week and different times of the day in order to maximize the chances of contacting a sample case. This type of calling schedule helps to keep the number of ring-no-answer cases to a minimum, which is important because a large number of "eligibility unknown" cases is a source of potential bias in the survey results.

In generating statistical estimates from the survey data, every sample case in every implemented sample wave must be accounted for, including every case for which a screening form was not returned and every case for which eligibility was not determined. Although these cases were classified as "eligibility status unknown," they cannot be ignored in the estimation process. Sampling weight calculations included adjustments for:

- That portion of the unknown-eligibility category of nonresponse estimated to be survey eligible, and
- The category of nonresponse due to failure of sample eligible to participate in the survey.

These sampling weight adjustments were made in an attempt to reduce the possible bias caused by missing information for sample cases. However, no adjustment can eliminate the potential for such bias. This can only occur when there are no cases for which eligibility status is unknown and no nonresponse.

4.4 ERRORS MADE IN IMPLEMENTING SURVEYS

Less than 5 percent of the 34,005 survey forms received during Year 3 for the eight Year 3 surveys and two Year 2 carryover surveys contained errors. The comparable error rate for forms received during Year 2 was 13 percent. The substantial drop in the error rate is a mark of success due to increased emphasis in training and diligent adherence to prescribed survey procedures.

Table 4-1 Comparison of Estimates of Survey Eligibles

	1980 Census			1987-88 Donnelley			
State	Number of Occupied Housing Units (1)	Percent Owner-Occupied (2)	Estimated Number of Owner-Occupied Single Family Housing Units with Telephone (3)*	Number of Housing Units with Telephone (4)	Estimated Number of Owner-Occupied Single Family Housing Units with Telephone (5)**	Ratio of Donnelley-Estimated Eligibles to Census-Estimated Eligibles (5) ÷ (3) = (6)	Ratio of Donnelley-Estimated Eligibles to Market Statistics' Estimated Eligibles (5) ÷ (9) = (7)
AK	131,463	58.3	69,883	69,588	38,136	.55	.37
GA	1,871,652	65.0	1,109,272	1,498,948	915,857	.83	.70
IA	1,053,033	71.8	689,392	875,846	591,126	.86	.81
ME	395,184	70.9	255,473	380,603	253,657	.99	.89
NM	441,466	68.1	274,122	309,660	198,226	.72	.59
OH	3,833,828	68.4	2,391,048	2,863,064	1,840,836	.77	.73
VT	178,325	68.7	111,704	184,031	118,844	1.06	.93
WV	686,311	73.6	460,573	471,593	326,267	.71	.66

	December 1986 Market Statistics			State Radon Survey Estimates				
State	Number of Occupied Housing Units (8)	Estimated Number of Owner-Occupied Single Family Housing Units with Telephone (9)*	Ratio of Market Statistics' Estimated Eligibles to Census Estimate Eligibles (9) ÷ (3) = (10)	Sample Sizes (11)	Estimated Number of Survey Eligibles Housing Units (12)	Ratio of Survey-Estimated Eligibles to Census Estimate Eligibles (12) ÷ (3) = (13)	Ratio of Survey-Estimated Eligibles to Market Statistics (12) ÷ (9) = (14)	Ratio of Survey-Estimated Eligibles to Donnelley Estimate Eligibles (12) ÷ (5) = (15)
AK	194,000	103,126	1.48	1,127	38,287	.55	.37	1.00
GA	2,220,100	1,315,787	1.19	1,534	826,452	.75	.63	.90
IA	1,116,200	730,745	1.06	1,381	593,815	.86	.81	1.00
ME	441,900	285,673	1.12	839	236,917	.93	.83	.93
NM	544,600	338,162	1.23	1,728	193,112	.70	.57	.97
OH	4,062,000	2,533,352	1.06	1,734	1,843,743	.77	.73	1.00
VT	204,300	127,975	1.15	710	117,523	1.05	.92	.99
WV	731,900	491,167	1.07	1,006	324,038	.70	.66	.99

* Assuming 94 percent of owner-occupied units are one unit structures (1983). Also assuming 97 percent of housing units have a telephone (1981).

** Assuming column (2) percent owner-occupied and that 94 percent of these are one unit structures.

Table 4-2 Disposition of Sample Cases

	AK	GA	IA	ME	NM	OH	VT	WV	R6	R7
Sample Waves Activated	1-7 21-70	1-7 21-116	1-6 21-72	1-4 21-57	– 21-100	1-12 21-119	1-6 21-54	1-6 21-59	– –	– –
Sample Waves Used in Analysis	1-6 21-70	1-4 21-37 39-116	1-6 21-72	1-4 21-57	– 21-100	1-12 21-100 102-119	1-6 21-54	1-6 21-59	– –	– –
C/S Forms Received	2,800	5,059	2,896	2,045	3,956	5,500	2,000	2,250	1,012	803
Case Used in Analysis	2,800	4,950	2,900	2,050	4,000	5,500	2,000	2,250	1,105	813
Status Eligibility Status, Code Canister Acceptance										
A1 Eligible, Accepted	1,301	2,095	1,561	990	2,021	2,339	876	1,138	781	686
A2 Eligible, Refused	43	196	129	72	105	278	88	147	166	52
C Eligibility unknown	286	819	381	332	478	1,343	452	354	0	0
O Not Eligible	801	1,541	615	509	948	1,054	437	430	65	65
D Not a Residence	369	408	210	142	404	486	147	181	0	0
Total	2,800	5,059	2,896	2,045	3,956	5,500	2,000	2,250	1,012	803
U Usable Readings	1,127	1,534	1,381	839	1,728	1,734	715	1,006	740	669
Response Rate $(A_1/(A_1 + A_2)$	96.8%	91.4%	92.4%	93.2%	95.1%	89.4%	90.9%	88.6%	82.5%	93.0%
Participation Rate $(U/A_1 + A_2)$	83.9%	68.1%	82.1%	79.0%	81.3%	66.3%	74.2%	78.3%	78.1%	90.7%

APPENDIX A

Installation Procedures

INSTALLATION PROCEDURES

1. EXTRACTING DATA FROM THE DISKETTE

The diskette you have received contains three files:

- **DATA.FIL** - a compressed version of the screening measurement data collected in one year of the EPA/State Residential Radon surveys.
- **EXTRACT.EXE** - an executable program to extract and store the expanded version of the survey data file on your hard disk. The extract program will run on any IBM-compatible personal computer using the MS-DOS operating system, Version 2.0 or higher.
- **READ_ME.1ST** - a copy of these instructions.

To expand the compressed file onto your hard disk, place the diskette in the appropriate drive and change to this drive. (For example, type **A:** then press the Enter key.) Run the program by typing the command **EXTRACT**, then press the Enter key. The program will ask where you want to store the expanded file. Respond by entering a full DOS pathname and filename to specify the drive, directory and name for the expanded file. For example, you may enter **C:\SURVEY\FILE1.DAT**. Note that the directory to which the file will be written (**C:\SURVEY**) must already exist on your hard disk. If the file (**FILE1.DAT**) already exists on the directory, you will be asked if you want to overwrite the file. Enter **Y** or **N**, as appropriate. The expanded file will be created under the filename and directory specified.

The program will ask if you want to extract specific State/Indian lands data from the survey data file. (Note: Read the file size considerations noted below before deciding how to extract the data.) To extract all of the data in the file, enter **A**. Enter **S** to extract only a subset of the data, rather than the entire file. You may select state codes from the list as instructed by the program. Note that the codes must be entered exactly as listed. After selecting the states, enter **1** to extract the file. If you make a mistake, enter **2** to re-enter the list of codes. You may enter **3** at any time to see the list of codes again, or **0** to exit the program.

2. SIZE CONSIDERATIONS

The entire expanded file for this diskette requires approximately 1.3 Megabytes of disk space. The expanded file is a standard DOS text file, with fixed-length records, one record for each house returning useable measurements. The expanded data file contains 99 ASCII text characters on each record, followed by carriage return and linefeed characters at the end of each line of text. A description of the layout of information on each record is included in the documentation for this diskette as Appendix B. The variable names listed there are the names used in EPA's analysis of the survey data.

The expanded file may be imported into a variety of DOS application programs for display and/or analysis. Most DOS applications can import DOS text files. Analysis of the data will require the use of an application program and a computer with sufficient memory available to handle a file of the required size. This should be considered when the Extract program is run. If data for all states on the disk are extracted into a single expanded file and your computer does not have additional extended or expanded memory beyond the now standard 640 Kilobytes of DOS memory, the large size of the expanded file may cause problems in many applications.

Another consideration is the number of lines (records) in the expanded file. While Excel for Windows can accommodate over 16,000 lines of data, many spreadsheet programs have a limit of approximately 8,000 lines. The entire expanded file exceeds 8,000 lines and an error will occur when importing the file into Lotus 123, for example, although sufficient memory may be available. <u>If these size problems are a concern for your program or computer, we recommend extracting the data for each state into a separate file.</u> The resulting expanded files for each state will be much smaller and problems due to size will be avoided.

3. ACCESSING DATA IN THE EXPANDED FILE

The expanded file is sorted by county within states, so that all records for a given county are grouped together in the file. For users without access to more powerful software, selected portions of the data may be viewed and printed using any word processing program that accepts DOS text files as input. For example, in version 5.0 of Wordperfect this is accomplished by the [Control-F5, 1, 2] keystroke sequence. Select a smaller font or use the landscape page orientation to print all 99 columns of data.

To conserve disk space, the expanded file does not include blank spaces between adjacent entries on a record, so a simple printout of the file as received may difficult to read. It is also difficult to analyze the data using a word processing program. DOS spreadsheet and database application programs may be used to reformat, graph and/or analyze the data.

The expanded file may be imported into a Lotus 123 spreadsheet, for example, using the [/File, Import, Text] keystroke sequence, if sufficient memory is available. The specific variables on each record may be parsed into individual numeric and label cells using the [/Data, Parse, Format, Create] keystroke sequence to specify the columns with the desired information. Then set the Input and Output ranges from the data parse menu, followed by Go. Other spreadsheet and database packages have specific procedures for importing DOS text file specified in the user reference manual.

4. CONSIDERATIONS FOR DATA ANALYSIS

This file reports short-term screening level radon measurements, conducted in accordance with prevailing EPA protocols in effect in the year of the survey. The file contains one record for each surveyed home with a useable radon measurement collected during the survey. Some data fields may have missing entries on certain records. Although attempts were made to gather complete information on each useable radon test, it was not possible to complete all items for all surveyed homes. Missing data items are indicated by a blank data field or by a single period in the data field.

The radon concentrations were estimated using a laboratory counting procedure on the exposed charcoal canisters, with a correction made for counts due to background radiation. This correction results in negative estimates of the radon concentration in some homes. These negative numbers should be considered a result of measurement error. In reality, radon concentrations are always non-negative.

The percent error variable recorded on the data file is the percentage measurement error reported by the EPA laboratory. This 2-sigma error bound was calculated based on the expected counting errors involved in the measurement process. No percentage measurement errors were reported by the laboratory for radon activities less than about 0.50 pCi/L. In the database the percent error variable is set to 0.0 on these records. For this variable, a percent error value of 0.0 should be treated as a missing value. In reality, the percentage measurement error associated with these measurements is very large.

The two problems noted above both derive from the lack of a specified Lower Limit of Detection (LLD) for the state survey data. One solution to both problems is to use the percent error variable to define the LLD for the radon activity variable. If the percent error is 0.0 and the radon activity is 0.5 pCi/L or less, then the radon activity measurement is below the LLD for the laboratory and its actual numeric value is meaningless. Alternatively, the negative activity values may be set to a small non-negative number, such as 0.05 pCi/L. This alternative method was used to calculate the survey statistics reported in this documentation.

APPENDIX B

Record Layout for State Residential Radon Surveys

Record Layout for State Residential Radon Surveys

Variable	Position	Type	Length	Description
STATE	1-2	A	2	State Postal Abbreviation (R5, R6, R7, RB, RC, RN are Indian Nations)
STATE2	3-4	A	2	State Postal Abbreviation for Indian Land Surveys (STATE = STATE2 for all other records)
STFIPS	5-6	N	2	State FIPS Code
ZIP	7-11	A	5	Zip Code
REGION	12-13	N	2	Analysis Region Code
TYPEBLDG	14	N	1	Type of Building 0 = unknown 1 = single family 2 = multi-family 3 = business 4 = school 5 = other
FLOOR	15	N	1	Floor Level 0 = basement 1 = first floor 2 = second floor or above 9 = unknown
ROOM	16	N	1	Type of Room 0 = unknown 1 = bedroom 2 = family room 3 = living room 4 = unfinished basement 5 = office 6 = classroom 7 = other

Record Layout for State Residential Radon Surveys - continued

Variable	Position	Type	Length	Description
BASEMENT	17	A	1	Is There a Basement in the Building? blank = unknown Y = Yes N = No
WINDOOR	18	A	1	House Closed or Open During Test blank = unknown O = Open C = Closed
REP	19-20	N	2	Replicate Number
STRATUM	21-22	N	2	Stratum Number
WAVE	23-25	N	3	Wave Number
STARTTM	26-29	N	4	Start Time of Test (HHMM)
STOPTM	30-33	N	4	Stop Time of Test (HHMM)
STARTDT	34-39	N	6	Start Date of Test (MMDDYY)
STOPDT	40-45	N	6	Stop Date of Test (MMDDYY)
ACTIVITY	46-53	N	8.1	Activity (pCi/L)
PCTERR	54-61	N	8.1	Percent Error (2-sigma)
ADJWT	62-74	N	13.6	Analysis Weight
DUPFLAG	75	N	1	Duplicate Flag 0 = activity from single canister 1 = average activity from duplicate canisters
ZIPFLAG	76	N	1	Flag for Zip Code (ZIP) 0 = believed accurate 1 = questionable

Record Layout for State Residential Radon Surveys - continued

Variable	Position	Type	Length	Description
CNTYFIPS	77-79	N	3	County FIPS Code
COUNTY	80-99	A	20	County Name

APPENDIX C

Description of Sample Allocation Used for Each State

ALASKA (02)

Allocation #2 was used.
Expected DEFF = 1.229

Stratum	Geographical Region	Estimated Population	Estimated Number of Telephones	Expected Number of Canisters	Relative Sampling Rates
1	Anchorage Borough	237,000	122,000	377	1X
2	Fairbanks and Surrounding Area	76,000	25,000	387	5X
3	Southeastern Alaska	58,000	20,000	309	5X
4	South Central Alaska	97,000	46,000	427	3X
	Total:	468,000	213,000	1,500	

GEORGIA (13)

Allocation #4 was used.
Expected DEFF = 1.223

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1A	GA01 (H), GA02 (M)	530	2.0 x
1B	GA01 (H), GA02 (M)	627	1.0 x
2A	GA01 (H), GA02 (M)	159	4.0 x
2B	GA02 (M)	327	2.0 x
2C	GA02 (M)	204	1.0 x
3A	GA02 (M)	336	4.0 x
3B	GA02 (M)	67	1.0 x
	Total:	2,250	

IOWA (19)

Allocation #3 was used.
Expected DEFF = 1.2646

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	IA01 (H)	153	3.0 x
2	IA01 (H)	183	3.0 x
3	IA01 (H)	159	3.0 x
4	IA01 (H)	156	3.0 x
5	IA01 (H)	193	2.0 x
6A	IA01 (H)	9	3.0 x
6B	IA01 (H)	41	2.0 x
6C	IA01 (H)	145	1.0 x
7A	IA01 (H)	70	2.0 x
7B	IA01 (H)	125	1.0 x
8	IA01 (H)	167	4.0 x
9A	IA01 (H)	32	3.0 x
9B	IA01 (H)	118	2.0 x
	Total:	1,550	

MAINE (23)

Allocation #4 was used.
Expected DEFF = 1.095

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1	ME01 (H)	254	1.0 x
2	ME02	137	1.0 x
3	ME03	119	1.5 x
4	ME04	123	1.5 x
5	ME05	105	2.0 x
6A	ME06	46	4.0 x
6B	ME06	83	1.0 x
7	ME07 (L)	101	2.0 x
	Total:	968	

NEW MEXICO (35)

Allocation #3 was used.
Expected DEFF = 1.199

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1A	NM01 (H), NM02 (M)	485	2.0 x
1B	NM01 (H), NM02 (M), NM03 (L)	609	1.0 x
2A	NM01 (H), NM02 (M), NM03 (L)	420	4.0 x
2B	NM01 (H)	98	1.0 x
3A	NM01 (H), NM02 (M)	36	4.0 x
3B	NM01 (H), NM02 (M)	15	2.0 x
3C	NM01 (H), NM03 (L)	173	1.0 x
4A	NM02 (M)	263	1.0 x
4B	NM03 (L)	51	2.0 x
	Total:	2,250	

OHIO (39)

Allocation #4 was used.
Expected DEFF = 1.157

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1A	OH01 (H), OH02 (M)	91	3.0 x
1B	OH01 (H), OH02 (M)	441	2.0 x
2A	OH02 (M)	62	3.0 x
2B	OH02 (M)	580	1.0 x
3A	OH01 (H)	49	3.0 x
3B	OH01 (H), OH02 (M)	164	2.0 x
3C	OH01 (H), OH02 (M)	282	1.0 x
4A	OH01 (H), OH02 (M)	92	3.0 x
4B	OH01 (H), OH02 (M)	488	2.0 x
	Total:	2,250	

VERMONT (50)

Allocation #3 was used.
Expected DEFF = 1.1576

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1A	VT01 (H)	9	2.5 x
1B	VT01 (H)	142	1.0 x
2	VT02 (M^+)	139	2.5 x
3	VT03 (M^-)	154	2.0 x
4	VT04 (L^+)	150	2.0 x
5	VT05 (L^-)	157	2.5 x
	Total:	750	

WEST VIRGINIA (54)

Allocation #2 was used.
Expected DEFF = 1.083

Stratum	Geological Classification Expected Radon Level	Canisters	Relative Sampling Rates
1A	WV01 (H)	115	2.0 x
1B	WV01 (H)	108	1.0 x
2A	WV01 (H)	178	2.0 x
2B	WV02 (M)	98	1.0 x
3A	WV01 (H)	86	2.0 x
3B	WV01 (H), WV02 (M)	539	1.0 x
	Total:	1,125	

Region 6 Indians (08, 35, 48, 49)

Code	State	Stratum	Canisters	Rate
(11)	NM	1	138	2.0 x
(19)	NM	2	5	1.0 x
(07)	NM		53	
(25)	NM		13	
(20)	NM		10	
(08)	NM		39	
(04)	NM		15	
			135	
(03)	NM	3	55	1.0 x
(10)	NM		81	
(05)	NM		38	
			174	
(26)	NM	4	148	2.0 x
(15)	NM		63	
			211	
(01)	CO	5	96	4.0 x
(02)	CO		103	
(28)	UT		38	
			237	
(23)	NM	6	62	1.0 x
(09)	NM		75	
(21)	NM		24	
(22)	NM		52	
(16)	NM		34	
(06)	NM		22	
(12)	NM		14	
(14)	NM		9	
(17)	NM		10	
(18)	NM		28	
(13)	NM		7	
(24)	NM		6	
(27)	TX		13	
			356	
		Total:	1,251	

Region 7 Indians (19, 31, 38, 46)

Code	State	Stratum	Canisters	Rate
(36-38)	SD	1	19	1.0 x
(40)	SD	2	17	1.0 x
(21-35)	SD	3	64	1.0 x
(41-55)	SD	4	85	1.0 x
(56)	SD	5	17	1.0 x
(57-72)	SD	6	92	1.0 x
(73-83)	SD	7	32	1.0 x
(89-92)	SD	8	26	1.0 x
(39)	SD	9	8	1.0 x
(84-88)	SD	10	41	1.0 x
(12-14)	ND	11	52	1.0 x
(08-11)	ND	12	52	1.0 x
(16-17)	ND	13	130	1.0 x
(15)	ND	14	9	1.0 x
(02-07)	ND	15	45	1.0 x
(18)	NE	16	22	1.0 x
(20)	NE	17	23	1.0 x
(19)	NE	18	99	1.0 x
(01)	IA	19	10	1.0 x
		Total:	753	

Table C-1 Distribution of Canisters per County for Alaska

COUNTY	REGION	# CANISTERS
ALEUTIANS EAST	0	0
ALEUTIANS WEST	0	0
ANCHORAGE	1	282
BETHEL	0	0
BRISTOL BAY	0	0
DILLINGHAM	0	0
FAIRBANKS NORTHSTAR	2	281
HAINES	3	12
JUNEAU	3	137
KENAI PENINSULA	4	135
KETCHIKAN GATEWAY	3	56
KODIAK ISLAND	4	27
LAKE AND PENINSULA	0	0
MATANUSKA-SUSITNA	4	60
NOME	0	0
NORTH SLOPE	0	0
NORTHWEST ARCTIC	0	0
PRINCE OF WALES-OUTER	3	0
SITKA	3	24
SKAGWAY-YAKUTAT-ANGOON	3	3
SOUTHEAST FAIRBANKS	2	31
VALDEZ-CORDOVA	4	31
WADE HAMPTON	0	0
WRANGELL-PETERSBURG	3	35
YUKON-KOYUKUK	2	13

Table C-1 Distribution of Canisters per County for Georgia

COUNTY	REGION	# CANISTERS
APPLING	3	9
ATKINSON	3	5
BACON	3	6
BAKER	2	0
BALDWIN	2	5
BANKS	1	6
BARROW	1	8
BARTOW	1	9
BEN HILL	2	4
BERRIEN	2	4
BIBB	2	23
BLECKLEY	2	7
BRANTLEY	3	4
BROOKS	2	2
BRYAN	3	9
BULLOCH	3	20
BURKE	2	5
BUTTS	1	6
CALHOUN	2	1
CAMDEN	3	15
CANDLER	3	2
CARROLL	1	25
CATOOSA	1	5
CHARLTON	3	5
CHATHAM	3	20
CHATTAHOOCHEE	2	0
CHATTOOGA	1	5
CHEROKEE	1	10
CLARKE	1	8
CLAY	2	1
CLAYTON	1	30
CLINCH	3	4
COBB	1	83
COFFEE	3	22
COLQUITT	2	17
COLUMBIA	2	24
COOK	2	1
COWETA	1	13
CRAWFORD	2	3
CRISP	2	8
DADE	1	7
DAWSON	1	2
DE KALB	1	76
DECATUR	2	9
DODGE	2	4

Table C-1 Distribution of Canisters per County for Georgia (Continued)

COUNTY	REGION	# CANISTERS
DOOLY	2	2
DOUGHERTY	2	11
DOUGLAS	1	26
EARLY	2	3
ECHOLS	2	1
EFFINGHAM	3	12
ELBERT	1	5
EMANUEL	2	9
EVANS	3	5
FANNIN	1	5
FAYETTE	1	31
FLOYD	1	18
FORSYTH	1	11
FRANKLIN	1	7
FULTON	1	75
GILMER	1	5
GLASCOCK	2	2
GLYNN	3	6
GORDON	1	7
GRADY	2	6
GREENE	1	3
GWINNETT	1	73
HABERSHAM	1	12
HALL	1	37
HANCOCK	2	6
HARALSON	1	9
HARRIS	2	2
HART	1	11
HEARD	1	1
HENRY	1	17
HOUSTON	2	18
IRWIN	2	1
JACKSON	1	6
JASPER	2	3
JEFF DAVIS	3	5
JEFFERSON	2	2
JENKINS	2	5
JOHNSON	2	4
JONES	2	6
LAMAR	1	3
LANIER	2	2
LAURENS	2	13
LEE	2	0
LIBERTY	3	12
LINCOLN	2	8

Table C-1 Distribution of Canisters per County for Georgia (Continued)

COUNTY	REGION	# CANISTERS
LONG	3	3
LOWNDES	2	10
LUMPKIN	1	1
MACON	2	1
MADISON	1	8
MARION	2	0
MCDUFFIE	2	18
MCINTOSH	3	2
MERIWETHER	1	9
MILLER	2	2
MITCHELL	2	8
MONROE	2	3
MONTGOMERY	2	3
MORGAN	1	6
MURRAY	1	9
MUSCOGEE	2	22
NEWTON	1	8
OCONEE	1	1
OGLETHORPE	1	4
PAULDING	1	4
PEACH	2	5
PICKENS	1	4
PIERCE	3	13
PIKE	1	6
POLK	1	8
PULASKI	2	3
PUTNAM	2	11
QUITMAN	2	0
RABUN	1	6
RANDOLPH	2	2
RICHMOND	2	26
ROCKDALE	1	24
SCHLEY	2	2
SCREVEN	2	5
SEMINOLE	2	1
SPALDING	1	19
STEPHENS	1	12
STEWART	2	1
SUMTER	2	12
TALBOT	2	3
TALIAFERRO	2	3
TATTNALL	3	7
TAYLOR	2	1
TELFAIR	2	5
TERRELL	2	2

Table C-1 Distribution of Canisters per County for Georgia (Continued)

COUNTY	REGION	# CANISTERS
THOMAS	2	11
TIFT	2	3
TOOMBS	3	20
TOWNS	1	5
TREUTLEN	2	3
TROUP	1	19
TURNER	2	3
TWIGGS	2	2
UNION	1	8
UPSON	1	11
WALKER	1	16
WALTON	1	13
WARE	3	22
WARREN	2	3
WASHINGTON	2	6
WAYNE	3	11
WEBSTER	2	3
WHEELER	2	1
WHITE	1	8
WHITFIELD	1	8
WILCOX	2	3
WILKES	2	9
WILKINSON	2	4
WORTH	2	1

Table C-1 Distribution of Canisters per County for Iowa

COUNTY	REGION	# CANISTERS
ADAIR	3	3
ADAMS	3	5
ALLAMAKEE	5	6
APPANOOSE	8	13
AUDUBON	2	6
BENTON	7	11
BLACK HAWK	5	55
BOONE	6	11
BREMER	5	17
BUCHANAN	5	14
BUENA VISTA	1	17
BUTLER	4	17
CALHOUN	2	5
CARROLL	2	17
CASS	3	14
CEDAR	7	6
CERRO GORDO	4	24
CHEROKEE	1	15
CHICKASAW	5	4
CLARKE	8	9
CLAY	1	9
CLAYTON	5	13
CLINTON	7	15
CRAWFORD	2	12
DALLAS	6	10
DAVIS	9	8
DECATUR	8	11
DELAWARE	5	9
DES MOINES	9	24
DICKINSON	1	12
DUBUQUE	5	38
EMMET	1	11
FAYETTE	5	11
FLOYD	4	12
FRANKLIN	4	17
FREMONT	3	5
GREENE	6	8
GRUNDY	6	6
GUTHRIE	2	10
HAMILTON	6	10
HANCOCK	4	8
HARDIN	6	4
HARRISON	2	12
HENRY	9	7
HOWARD	5	10

Table C-1 Distribution of Canisters per County for Iowa (Continued)

COUNTY	REGION	# CANISTERS
HUMBOLDT	4	9
IDA	2	3
IOWA	7	8
JACKSON	7	9
JASPER	6	13
JEFFERSON	9	13
JOHNSON	7	16
JONES	7	8
KEOKUK	9	7
KOSSUTH	4	19
LEE	9	17
LINN	7	41
LOUISA	9	8
LUCAS	8	8
LYON	1	11
MADISON	8	11
MAHASKA	8	12
MARION	8	33
MARSHALL	6	10
MILLS	3	12
MITCHELL	4	6
MONONA	2	10
MONROE	8	11
MONTGOMERY	3	7
MUSCATINE	9	20
O'BRIEN	1	13
OSCEOLA	1	4
PAGE	3	17
PALO ALTO	1	10
PLYMOUTH	1	15
POCAHONTAS	1	6
POLK	6	77
POTTAWATTAMIE	3	45
POWESHIEK	7	12
RINGGOLD	3	3
SAC	2	9
SCOTT	7	35
SHELBY	2	14
SIOUX	1	21
STORY	6	19
TAMA	7	8
TAYLOR	3	6
UNION	3	11
VAN BUREN	9	7
WAPELLO	9	15

Table C-1 Distribution of Canisters per County for Iowa (Continued)

COUNTY	REGION	# CANISTERS
WARREN	8	29
WASHINGTON	9	7
WAYNE	8	6
WEBSTER	6	11
WINNEBAGO	4	10
WINNESHIEK	5	10
WOODBURY	2	62
WORTH	4	6
WRIGHT	4	10

Table C-1 Distribution of Canisters per County for Maine

COUNTY	REGION	# CANISTERS
ANDROSCOGGIN	2	47
AROOSTOOK	7	102
CUMBERLAND	1	132
FRANKLIN	3	22
HANCOCK	5	53
KENNEBEC	2	61
KNOX	4	30
LINCOLN	4	18
OXFORD	3	42
PENOBSCOT	6	79
PISCATAQUIS	6	42
SAGADAHOC	4	34
SOMERSET	3	31
WALDO	4	27
WASHINGTON	5	40
YORK	1	79

Table C-1 Distribution of Canisters per County for New Mexico

COUNTY	REGION	# CANISTERS
BERNALILLO	1	406
CATRON	3	16
CHAVES	4	52
CIBOLA	1	6
COLFAX	2	91
CURRY	4	47
DE BACA	4	12
DONA ANA	3	86
EDDY	4	51
GRANT	3	60
GUADALUPE	4	8
HARDING	2	12
HIDALGO	3	18
LEA	4	50
LINCOLN	4	18
LOS ALAMOS	1	42
LUNA	3	49
MCKINLEY	1	53
MORA	2	17
OTERO	3	46
QUAY	4	10
RIO ARRIBA	2	72
ROOSEVELT	4	44
SAN JUAN	1	196
SAN MIGUEL	2	78
SANDOVAL	1	76
SANTA FE	2	73
SIERRA	3	41
SOCORRO	1	41
TAOS	2	47
TORRANCE	1	10
UNION	2	32
VALENCIA	1	25

Table C-1 Distribution of Canisters per County for Ohio

COUNTY	REGION	# CANISTERS
ADAMS	3	10
ALLEN	1	28
ASHLAND	2	20
ASHTABULA	2	15
ATHENS	4	14
AUGLAIZE	1	10
BELMONT	4	12
BROWN	3	5
BUTLER	3	33
CARROLL	2	7
CHAMPAIGN	3	12
CLARK	3	15
CLERMONT	3	12
CLINTON	3	9
COLUMBIANA	2	13
COSHOCTON	2	18
CRAWFORD	1	14
CUYAHOGA	2	120
DARKE	3	15
DEFIANCE	1	8
DELAWARE	1	20
ERIE	1	19
FAIRFIELD	4	31
FAYETTE	3	6
FRANKLIN	4	170
FULTON	1	6
GALLIA	4	11
GEAUGA	2	6
GREENE	3	25
GUERNSEY	4	13
HAMILTON	3	90
HANCOCK	1	15
HARDIN	1	17
HARRISON	2	7
HENRY	1	16
HIGHLAND	3	8
HOCKING	4	9
HOLMES	2	9
HURON	1	14
JACKSON	4	15
JEFFERSON	2	7
KNOX	1	14
LAKE	2	28
LAWRENCE	4	9
LICKING	4	29

Table C-1 Distribution of Canisters per County for Ohio (Continued)

COUNTY	REGION	# CANISTERS
LOGAN	3	19
LORAIN	2	21
LUCAS	1	71
MADISON	3	10
MAHONING	2	20
MARION	1	17
MEDINA	2	9
MEIGS	4	9
MERCER	1	12
MIAMI	3	22
MONROE	4	6
MONTGOMERY	3	67
MORGAN	4	2
MORROW	1	8
MUSKINGUM	4	24
NOBLE	4	6
OTTAWA	1	9
PAULDING	1	8
PERRY	4	12
PICKAWAY	4	7
PIKE	4	8
PORTAGE	2	6
PREBLE	3	4
PUTNAM	1	18
RICHLAND	1	29
ROSS	4	10
SANDUSKY	1	11
SCIOTO	4	13
SENECA	1	21
SHELBY	3	9
STARK	2	50
SUMMIT	2	60
TRUMBULL	2	34
TUSCARAWAS	2	13
UNION	1	6
VAN WERT	1	18
VINTON	4	2
WARREN	3	15
WASHINGTON	4	16
WAYNE	2	12
WILLIAMS	1	8
WOOD	1	18
WYANDOT	1	10

Table C-1 Distribution of Canisters per County for Vermont

COUNTY	REGION	# CANISTERS
ADDISON	1	26
BENNINGTON	4	58
CALEDONIA	5	51
CHITTENDEN	1	102
ESSEX	5	14
FRANKLIN	1	24
GRAND ISLE	1	12
LAMOILLE	2	29
ORANGE	5	43
ORLEANS	5	50
RUTLAND	4	70
WASHINGTON	2	101
WINDHAM	3	51
WINDSOR	3	79

Table C-1 Distribution of Canisters per County for West Virginia

COUNTY	REGION	# CANISTERS
BARBOUR	3	21
BERKELEY	1	19
BOONE	3	15
BRAXTON	3	15
BROOKE	2	25
CABELL	3	43
CALHOUN	3	3
CLAY	3	7
DODDRIDGE	3	0
FAYETTE	3	24
GILMER	3	8
GRANT	1	10
GREENBRIER	1	18
HAMPSHIRE	1	12
HANCOCK	2	34
HARDY	1	9
HARRISON	3	37
JACKSON	3	11
JEFFERSON	1	13
KANAWHA	3	108
LEWIS	3	15
LINCOLN	3	11
LOGAN	3	16
MARION	2	36
MARSHALL	2	18
MASON	3	7
MCDOWELL	3	8
MERCER	1	20
MINERAL	1	15
MINGO	3	10
MONONGALIA	2	20
MONROE	1	20
MORGAN	1	13
NICHOLAS	3	16
OHIO	2	47
PENDLETON	1	8
PLEASANTS	3	6
POCAHONTAS	1	18
PRESTON	2	31
PUTNAM	3	20
RALEIGH	3	38
RANDOLPH	1	25
RITCHIE	3	8
ROANE	3	6
SUMMERS	1	11
TAYLOR	2	11
TUCKER	2	6

Table C-1 Distribution of Canisters per County for West Virginia (Continued)

COUNTY	REGION	# CANISTERS
TYLER	3	8
UPSHUR	3	7
WAYNE	3	16
WEBSTER	3	5
WETZEL	2	12
WIRT	3	3
WOOD	3	44
WYOMING	3	19

APPENDIX D

Regional Radon Coordinators and Sources of Information Concerning Other State-Wide Radon Studies

Regional Radon Coordinators

EPA REGION	REGIONAL OFFICE	CONTACT
1	U.S. Environmental Protection Agency John F. Kennedy Federal Building Room 2311 Boston, MA 02203	Mona Haywood (617) 565-9402
2	U.S. Environmental Protection Agency 26 Federal Plaza Room 1137-L New York, NY 10278	Lorainne Koehler (212) 264-0546
3	U.S. Environmental Protection Agency (3AM12) 841 Chestnut Street Philadelphia, PA 19107	Lewis Felleisen (215) 597-8326
4	U.S. Environmental Protection Agency 345 Courtland Street, NE Atlanta, GA 30365	Paul Wagner (404) 347-3907
5	U.S. Environmental Protection Agency Mail Code (AT-18J) 77 West Jackson Blvd. Chicago, IL 60604	Julie Beckman (312) 886-6063
6	U.S. Environmental Protection Agency Air Enforcement Branch (6T-E) 1445 Ross Avenue Dallas, TX 75202	Michael Miller (214) 655-7550
7	U.S. Environmental Protection Agency 726 Minnesota Avenue Kansas City, KS 66101	Bob Hunt (913) 551-7611
8	U.S. Environmental Protection Agency (8HWM-RP) Suite 500 999 18th Street Denver, CO 80202	Milton W. Lammering (303) 293-1440
9	U.S. Environmental Protection Agency (A1-1) 75 Hawthorne Street San Francisco, CA 94105	Louise Hill (415) 744-1046
10	U.S. Environmental Protection Agency (AT-082) 1200 Sixth Avenue Seattle, WA 98101	Misha Vakoc (206) 553-7299

Sources of Information Concerning Other State-Wide Radon Studies		
STATE	AGENCY	CONTACT
New Jersey	Department of Environmental Protection 729 Alexander Road Princeton, NJ 08540	Robert Stern (800) 648-0394 (609) 987-6402
New York	State Health Department Bureau of Environmental Radiation Protection Corning Tower Albany, NY 12237	Laurence Keefe (800) 458-1158 (518) 458-6450
North Carolina	Department of Human Resources Radiation Protection Section 701 Barbour Drive Raleigh, NC 27603-2008	Dr. Felix Fong (919) 733-4283
Idaho	Department of Health and Welfare Bureau of Preventive Medicine 450 West State Street Boise, ID 83720	Janne Mitten (208) 334-5927
Florida	Department of Health and Rehabilitative Services 1317 Winewood Boulevard Tallahassee, FL 32399-0700	N. Michael Gilly (800) 543-8279 (904) 488-1525
South Carolina	Department of Health and Environmental Control Bureau of Radiological Health 2600 Bull Street Colombia, SC 29201	Nolan Bivens (803) 734-4700
Oregon	Department of Human Services Health Division 1400 SW 5th Avenue Portland, OR 97201	Ray Paris (503) 229-5797
Washington	Department of Health Office of Radiation Protection Airdustrial Building 5, LE-13 Olympia, WA 98504	Robert Mooney (206) 586-3303

STATE	AGENCY	CONTACT
Montana	Department of Health and Environmental Sciences Cogswell Building Helena, MT 59620	Adrian Howe (406) 444-3671
New Hampshire	Division of Public Health Serv. Bureau of Radiological Health 6 Hazen Drive Concord, NH 03301	Joy Hanington (603) 271-4674
Virginia	Department of Health Bureau of Radiological Health 109 Governor Street Richmond, VA 23219	Leslie Foldesi (800) 468-0138 (804) 786-5932
Nevada	Department of Human Resources Radiological Health Section 505 East King Street, Rm. 203 Carson City, NV 89710	Stan Marshall (702) 885-5394
Louisiana	Louisiana Nuclear Energy Division Department of Environmental Qual. P.O. Box 14690 Baton Rouge, LA 70898	Jay Mason (504) 925-4518

APPENDIX E

Procedures for Estimating Weighted Means, Proportions, Standard Errors, and Confidence Intervals for Indian Lands

Procedures for Estimating Weighted Means, Proportions,
Standard Errors and Confidence Intervals for Indian Lands

The EPA's Region 6 Indian lands consist of 28 reservations and EPA's Region 7 Indian lands consist of 19 reservations. For purposes of the radon survey, these areas were stratified according to reservation and a simple random sample of households was selected within each reservation or stratum. Formulas for generating estimates of weighted means, proportions and standard errors for the Indian land surveys are given below. An approximate 95 percent confidence interval can be derived by adding to and subtracting from the estimate two standard errors of the estimate.

NOTATION

Let, Y_{hi} = observed radon measurement for the i^{th} household in stratum h ($i=1, .., n_h$ and $h=1,...,H$);

W_{hi} = sampling weight associated with Y_{hi} ;

$$J_{rh} = \begin{cases} 1 & \text{if stratum } h \text{ is included in the } r^{th} \text{ region} \\ 0 & \text{otherwise;} \end{cases}$$

$$I_{xhi} = \begin{cases} 1 & \text{if measurement on } i^{th} \text{ household in stratum } h \text{ is greater than X pCi/L} \\ 0 & \text{otherwise;} \end{cases}$$

n_h = number of sample households in stratum h;

H = number of strata;

$$N_h = \sum_{i=1}^{n_h} W_{hi} \quad ;$$

$$N_r = \sum_{h=1}^{H} J_{rh} \quad N_h \quad ;$$

$$N = \sum_{h=1}^{H} N_h \quad ;$$

$$s^2_h = \frac{\sum_{i=1}^{n_h} Y^2_{hi} - \left| \sum_{i=1}^{n_h} Y_{hi} \right|^2 / n_h}{n_h - 1} \; ; \quad \text{and}$$

$$S.E.(est.) = [Var(est.)]^{1/2} .$$

ESTIMATION:

The true mean radon level for the h^{th} stratum or reservation can be estimated as

$$Y^*_h = \frac{\sum_{i=1}^{n_h} W_{hi} \, Y_{hi}}{\sum_{i=1}^{n_h} W_{hi}} \tag{1}$$

The mean radon level for the r^{th} region, consisting of two or more reservations, is given by the weighted average of the strata making up the region, namely

$$Y^*_r = \frac{\sum_{h=1}^{H} J_{rh} \, N_h \, Y^*_h}{N_r} \tag{2}$$

The variance of Y^*_r is estimated as

$$Var(Y^*_r) = \frac{1}{N^2_r} \sum_{h=1}^{H} J_{rh} \, N_h(N_h - n_h) \left| \frac{s^2_h}{n_h} \right| , \tag{3}$$

and the standard error is obtained as $S.E. (Y^*_r) = [Var(Y^*_r)]^{1/2}$. A weighted average of all strata means provides an estimate of the overall mean,

$$Y^* = \frac{\sum_{h=1}^{H} N_h \, Y^*_h}{N} \tag{4}$$

The variance of Y^* is estimated as

$$Var(Y^*) = \frac{1}{N^2} \sum_{h=1}^{H} N_h(N_h - n_h) \left| \frac{S^2_h}{n_h} \right| , \tag{5}$$

and the standard error is obtained as $S.E.(Y^*) = [Var(Y^*)]^{1/2}$.

The true proportion of households in the h^{th} stratum with radon levels exceeding X pCi/L can be estimated as

$$P^*_h = \frac{\sum_{h=1}^{H} W_{hi} \, I_{xhi}}{\sum_{i=1}^{n_h} W_{hi}} \tag{6}$$

The proportion of households in the r^{th} region (i.e., combination of reservations) with radon levels exceeding X pCi/L is given by

$$P^*_r = \frac{\sum_{h=1}^{H} J_{rh} \, N_h \, P^*_h}{N_r} \tag{7}$$

The variance of P^*_r is estimated as

$$Var(P^*_r) = \frac{1}{N^2_r} \sum_{h=1}^{H} J_{rh} \, N_h(N_h - n_h) \left| \frac{P^*_h(1-P^*_h)}{n_h-1} \right| , \tag{8}$$

and the standard error is obtained as $S.E.(P^*_r) = [Var(P^*_r)]^{1/2}$. A weighted average of all strata proportion provides an overall proportion, namely

$$P^* = \frac{\sum_{h=1}^{H} N_h P^*_h}{N} . \tag{9}$$

The estimated variance of P^* is given by

$$Var(P^*) = \frac{1}{N^2} \sum_{h=1}^{H} N_h(N_h - n_h) \left[\frac{P^*_h(1-P^*_h)}{n_h - 1} \right] , \tag{10}$$

and the standard error is obtained as $S.E.(P^*) = [Var(P^*)]^{1/2}$.